Rapid Silicon Dioxide Film Formation on Clean Silicon Surfaces

by

Gene Andrew Danko

A dissertation submitted to the Johns Hopkins University

in conformity with the requirements for the degree of

Doctor of Philosophy

Baltimore, Maryland

1992

FINAL REPORT

N00014-89-J-1265

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	19 Oct. 1992	FINAL 1 Dec. 88 - 30 Jun 91

4. TITLE AND SUBTITLE
The Influence of the Kinetics of Formation on the Properties of SiO_2 Films on Si
See DOCUMENT

5. FUNDING NUMBERS
Grant No: N00014-89-J-1265
Modification No: P00001
R&T Proj: go14002---02
ACO Code: N66002
ONR Code: 1114SS
CACE Code: 5L406

6. AUTHOR(S)
Jerome Kruger and Gene Danko

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Johns Hopkins University
Department of Materials Science & Engineering
102 Maryland Hall, 3400 North Charles Street
Baltimore, Maryland 21218

8. PERFORMING ORGANIZATION REPORT NUMBER
G.42.5033

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Department of the Navy
Office of Naval Research
ATTN: Code 1513:FF
800 North Quincy Street
Arlington, Virginia 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Future generation of silicon-based microelectronic circuits will require ever-smaller devices, new classes of devices, and demands for higher reliability, thereby requiring further refinements of silicon planar technology. An understanding of the kinetics of film formation and optical properties of ultrathin silicon dioxide films on a parent silicon substrate is necessary to measure and predict the behavior of such devices.

A high-speed ellipsometer and growth chamber were constructed to measure the growth rate of SiO_2 on hot silicon substrates from which the prior native oxide had been removed.

Data gathered from temperatures between 800° C and 1000° C for three substrate orientations ((100), (111), and (110)) reveal the dependence of the refractive index of SiO_2 as a function of oxide thickness. No orientation effects were found. Kinetic measurements reveal two new linear growth regions with activation energies of E_I = 0/603 eV and E_{II} = 0.794 eV, respectively. X-ray photoelectron spectroscopy provides chemical evidence of oxygen supersaturation and a coesite-like structure near the oxide-substrate interface.

The results will provide baseline data necessary for radiation hardening assessments, data to aid the development of the next generation of ellipsometric thin film standards, and will permit process designers to develop thinner device oxides. The instrumentation developed for this work may have commercial applicability for process control feedback and *in situ* quality assurance.

14. SUBJECT TERMS
SiO_2 Films, Si, Ellipsometry, X-ray Photoelectron Spectroscopy
Kinetics of Film Formation, High Temperature Oxidation

15. NUMBER OF PAGES
168

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

Rapid Silicon Dioxide Film Formation on Clean Silicon Surfaces

by

Gene Andrew Danko

Abstract

Future generations of silicon-based microelectronic circuits will require ever-smaller devices, new classes of devices, and demands for higher reliability, thereby requiring further refinements of silicon planar technology. An understanding of the kinetics of film formation and optical properties of ultrathin silicon dioxide films on a parent silicon substrate is necessary to measure and predict the behavior of such devices.

A high-speed ellipsometer and growth chamber were constructed to measure the growth rate of SiO_2 on hot silicon substrates from which the prior native oxide had been removed.

Data gathered from temperatures between 800° C and 1000° C for three substrate orientations ((100), (111), and (110)) reveal the dependence of the refractive index of SiO_2 as a function of oxide thickness. No orientation effects were found. Kinetic measurements reveal two new linear growth regions with activation energies of $E_I = 0.603$ eV and $E_{II} = 0.794$ eV, respectively. X-ray photoelectron spectroscopy provides chemical evidence of oxygen supersaturation and a coesite-like structure near the oxide-substrate interface.

The results will provide baseline data necessary for radiation hardening assessments, data to aid the development of the next generation of ellipsometric thin film standards, and will permit process designers to develop thinner device oxides. The instrumentation developed for this work may have commercial applicability for process control feedback and *in situ* quality assurance.

Acknowledgments

This dissertation, which bears the name of one author, is actually the product of many minds. Foremost among these is Professor Jerome Kruger. Jerry's guidance and scientific insight has imbued me with talents which will last for a lifetime. It has been a high honor to study under one of the foremost members of the international community of corrosion science.

I wish to express my gratitude to Dr. Akos Revesz, whose name the reader will see on the technical papers within. Financial support was provided through the Office of Naval Research as overseen by Dr. Alvin Goodman, who had enough patience to let the project proceed at a natural pace. Thank you, gentlemen, for allowing me to learn along the way.

Within the Department of Materials Science and Engineering I must thank students, faculty, and staff. Of special mention are the department machinists, Mike Franckowiak and Walt Krug, upon whom I could rely for design assistance, work of superior quality, and access to tools and techniques without which this project could not have succeeded.

At the National Institute of Standards and Technology, Dr. Bernard Hockey, Dr. Lawrence Cook, Dr. Richard Ricker, Ms. Jonice Harris, and Mr. Art Sessoms of the Materials Science and Engineering Laboratory provided access to specialized equipment that was unavailable at the university. Within the Semiconductor Electronics Division, Dr. Deane Chandler-Horowitz and Dr. James Ehrstein provided independent ellipsometry and resistivity measurements of my materials. Their help is greatly appreciated.

And, of course, thanks to family (Mom) and friends (like Jim, who kept my car running) who brought me balance and kept me sane. This work is dedicated to all of those who have touched my life.

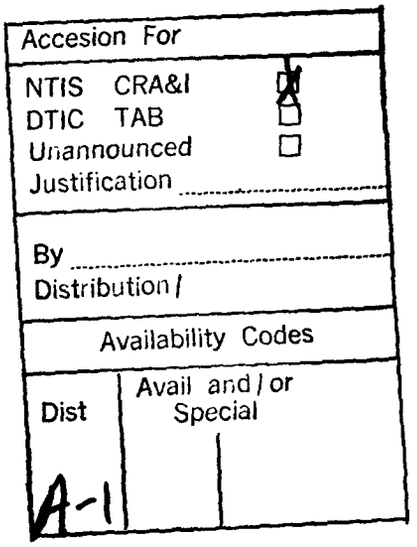


Table of Contents

List of Tables

List of Illustrations

Section I

Introduction and Background

Planar fabrication techniques are still the most widely used methods for the production of silicon-based microelectronics devices. Rapid advances over the last 50 years have evolved the technology to its present state, and with suitable modifications the silicon planar process should continue to dominate well into the next century.

The silicon planar process utilizes a repeated growth/etch schedule of oxides on a silicon substrate. The oxide layers serve as diffusion masks over which selected dopants are applied, which are driven into the silicon to modify its electronic characteristics. Successive applications of different dopants and different masks are performed to create a functional solid-state device. The growth and interconnection of devices on one piece of silicon results in a monolithic device, commonly referred to as an integrated circuit.

The planar process relies on the interrelationships of the properties of silicon and its oxide, SiO_2. The structural perfection and adhesion of the thermally grown oxide on its silicon substrate are superior to other semiconductor/oxide systems e. g. germanium or gallium arsenide. This distinction is further enhanced by the very different chemical behaviors of oxide and substrate; differences in diffusion profiles and susceptibility to environmental attack endow this system with the advantage of selectivity in processing. It is this fact that has been so cleverly exploited to produce the microelectronic devices of today.

Muller and Kamins (1) penned the eloquent statement:

> "Successful engineering rests on two foundations. One is a mastery of underlying physical concepts; a second foundation, at least of equal importance, is a perfected technology — a means to translate engineering concepts into useful structures."

It is this spirit that has guided the field of semiconductor technology. As the dimensions of silicon-based microelectronic devices continue to shrink, accurate thickness control is important for the fabrication of metal-oxide semiconductors (MOS). Oxide gate thicknesses must shrink as performance demands increase. Dielectric behavior of these oxides must be understood to predict device performance, which from the manufacturer's standpoint means that performance requirements drive the designer, who then relies on the researcher to provide adequate growth rate data for the proposed process. The *kinetics* of the silicon oxidation reaction need to be understood all the way back to nucleation at the bare surface as the needed oxide films get thinner and thinner. Quantum tunneling devices are now being explored which require oxides only 3 nanometers thick.

In addition to thickness control, *thickness determination* of SiO_2 films is becoming increasingly important. Since SiO_2 is primarily used as a mask to control dopant deposition on the underlying silicon substrate, exact knowledge of film thickness is helpful to properly design the etching conditions by which the mask is constructed. Most wet-etch techniques (e. g. solutions of hydrofluoric acid) are highly specific to the oxide but attack isotropically, resulting in undercutting of the mask. This phenomenon is only partially controllable through process manipulation, and the end result is a minimum device size beyond which production yields fall. Dry-etch techniques can be designed that are anisotropic (such as plasma etching), however, their usefulness is limited by a relative lack of selectivity which can result in damage to the substrate.

Semiconductor devices are susceptible to radiation damage. High-energy radiations — electrons, protons, gamma rays, neutrons — can produce point defects in the crystal lattice of the device which in turn agglomerate to produce donor sites, acceptor sites, or recombination centers. Such behavior always degrades the performance of the device.

Under normal conditions, defects anneal out readily with some time constant, τ. For hostile applications such as space flight or in nuclear reactors, radiation hardened devices can be designed. Commonly, layers of polysilicon and/or SiO_2 are used to protect the integrated circuitry by serving as sinks for the radiation damage. Assessment of both total radiation damage and damage rate can be performed on these layers by observing changes in their optical behavior. Success requires baseline data on the optical properties of thin SiO_2 layers, which has been the subject of much controversy in the literature. Establishment of the *film thickness/optical properties relationship* is a must as device complexity and reliability requirements advance.

Optical Properties

Archer (2,3) performed a series of ellipsometric measurements that confirmed a match between experimental and theoretical data in the Si/SiO_2 system. This work was important in establishing ellipsometry as a useful technique for semiconductor characterization. Goodman and Breece (4) measured film thicknesses on freshly cleaned wafers before and after short-term exposure to dry and wet oxygen at 600° C. While reporting reproducible production of thin (~3 nm) oxides, they admitted to disparities in the assumed optical constants of silicon. Nevertheless, their observations are accepted as sound.

An important experiment was reported by Taft and Cordes in 1979 (5) . A thick (200 nm) SiO_2 film was exposed to hydrofluoric acid and periodically measured ellipsometrically. This etchback experiment should have yielded decreases in film thickness but no change in the SiO_2 refractive index. Surprisingly, the refractive index did indeed change. Corrections for thermally-induced stress birefringence (generated on wafer cooling, due to

the mismatch in thermal expansion coefficients) were insufficient to explain the behavior; the insertion of an "interlayer" 0.6 nm thick with a refractive index of n = 2.8 into numerical calculations provided a solution that fit the data. Aspnes and Theeten (6, 7) and Theeten and Aspnes (8) confirmed this finding with a series of spectroscopic ellipsometric measurements and theoretical analyses.

To date, no hard evidence exists that proves or disproves the physical existence of the interlayer. Candela *et al.* (9) assume an interlayer in the certification of NIST SRM-2530, ellipsometric parameters Δ and Ψ and derived thickness and refractive index of a silicon dioxide layer on silicon. Chandler-Horowitz (10) has found that the interlayer model results in better mathematical solutions for the standard films, but that no physical structure has been identified to account for this behavior.

Kinetics

Deal and Grove (11) presented the first detailed model of silicon oxidation kinetics. Their linear-parabolic rate law identified two growth regimes which follow different kinetic equations: the linear "thin film" region, and the parabolic "thick film" region. Their argument for thick-film (>15 nanometers) growth was well developed and is still accepted as fundamentally sound. The thin film regime was explained by Deal and Grove in terms of space-charge assisted oxygen transport; under conditions of dry oxidation, an extrinsic Debye length of L_D =15 nanometers was calculated based on the equilibrium solubility of oxygen in SiO_2. When oxidation was carried out in the presence of water vapor, no thin film rate enhancement was observed, in good agreement with a calculated extrinsic Debye

length of only L_D =0.6 nanometers. Controversy surrounding this initial phase of growth has led to several competing paradigms for thin film growth.

Models for Thin Film Kinetics — These paradigms fall into three broad categories (see Massoud, Plummer, and Irene (12)): space-charge effects, oxide-structure effects, and oxide-stress effects. We shall consider each of these in turn.

In 1967 Grove (13) proposed a coupled diffusion mechanism (after Jorgensen (14)) where adsorbed O_2 would enter the oxide and pick up an electron, thereby forming an ion-hole pair. The highly mobile hole would diffuse toward the Si-SiO_2 interface, dragging the superoxide ion behind it. Hu(15) proposed an oxygen chemisorption step in which O^- ions could be formed at the gas-oxide interface. The low mass atomic ion could tunnel across thin oxides (<15 nanometers) and react quickly with silicon at the oxide-substrate interface. Hamasaki (16) proposed a mechanism in which interfacial oxidation results in a positively charged oxide. In the early stages of film growth, the charge is negligible and the negatively charged oxidant can flow inward. As the oxide thickens, the positive charge increases to set up a counter-field which suppresses inward motion of the oxidant.

Revesz and Evans (17) introduced the concept of microchannels (after Ing et al (18)) through the oxide, which would provide a conduit for the transport of molecular oxygen to the silicon-SiO_2 interface. This mechanism would work in parallel with the diffusion of O_2 through the SiO_2 film, but would become increasingly ineffective as the oxide thickened due to the difficulty in maintaining percolation paths through the random network of microchannels (19). The observed lack of rate enhancement for wet oxidation could be explained by the greatly increased SiOH defect density, which would disrupt channel

formation. Hopper, Clarke, and Young (20), in an ellipsometric study of silicon oxidation, recognized the need for more than one set of parameters to fit their experimental data. They stated that channel formation was one possibility, or the presence of ionic and molecular species simultaneously undergoing transport through the oxide (after the work of Jorgensen(21) and Raleigh (22)). Gibson and Dong (23) report pores in a TEM study of 9 nanometer SiO_2 films.

Several investigators have examined the role of stress within the oxide layer. Thermal stresses develop from differences in thermal expansion coefficients; these stresses appear on cooling of the substrate, leaving the oxide in a state of compressive stress, but at growth temperatures these thermal stresses do not exist. Intrinsic stresses result from the molar volume mismatch between the silicon substrate and the SiO_2 film. These stresses do develop at growth temperatures. Borden (24) found that intrinsic stress was a function of temperature: compressive for SiO_2 below 975° C, negligible from 975° C to 1000° C, and slightly tensile at higher temperatures. Investigators at this point take two separate tacks: EerNisse (25) and Massoud *et al.* (26) viewed viscous flow in the silicon as responsible for enhanced oxidation (e.g. due to dislocation generation). Fargeix *et al.* (27) suggested a slowing of the interstitial diffusion mechanism due to developing compressive stress in the oxide near the interface.

Therefore, the literature shows that there are wide disparities in the understanding of silicon oxidation mechanisms. Investigators have established that some different mechanism is operative in the very early stages of film formation, and that this (these) mechanism(s) works in parallel with the molecular oxygen diffusion mechanism first modeled by Deal and Grove.

Current Understanding — Most recently, Han and Helms (28, 29) and Revesz *et al.* (30) have employed ^{18}O tracer studies to localize incorporation sites of oxygen in the SiO_2 structure. Both studies postulate a parallel mechanism to molecular O_2 diffusion operative in the thin film limit; Han and Helms (28) results are in agreement with those of Hopper *et al.* (20), Massoud *et al.* (12), and Revesz *et al.* (19) although details of the approximations involved vary. Han and Helms report a high concentration of ^{18}O not only at the Si-SiO_2 interface, but also at the SiO_2-gas interface, and a low concentration throughout the bulk. Their data imply a parallel oxygen transport mechanism which favors the outward migration of oxygen vacancies generated by the injection of unreacted (or partially reacted) silicon atoms into the oxide. As these atoms react within the oxide, oxygen is migrating inward, leading to a concentration of oxygen vacancies at the surface. A surface reaction is thus predicted, which may account for the fast initial oxidation rate.

Revesz *et al.* (21) describe in some detail a mechanism of molecular O_2 diffusion to account for the distribution of ^{18}O in Han and Helms' tracer study. They claim that oxygen atoms from interstitially dissolved O_2 can exchange with SiO_2 network atoms (so-called "interstitialcy-diffusion"), leading to a distribution of tracer atoms in the bulk oxide while allowing the normal O_2 interfacial reaction to take place. No surface reaction is thought to occur, for no unreacted silicon ever reaches the surface.

Oxide-stress effects and stress relaxation have been addressed by several investigators. Doremus (31) has developed a model which permits stress relaxation in films formed above 950° C. At 1200° C the annealing is so rapid that no strain remains after film growth. Doremus' model states that deviation from the linear-parabolic rate law below 950° C is due to unrelaxed strains in the oxide; these strains are also responsible for the larger index of refraction (n = 1.472 for 800° C formed film versus n = 1.460 for vitreous SiO_2). Taft

(32) postulates that interfacial shear stresses account for the differences in index of refraction, are localized to the first few molecular distances from the interface (and thus are essentially planar), and can be relieved by viscous flow perpendicular to the interface. Mrstik *et al.* (33) refute viscoelastic stress relief in favor of a structural reordering away from the interface. They state that SiO_2 growth is quasi-epitaxial with the underlying silicon, and that stress relief is achieved through polymorphic transformation within the oxide away from the interface.

Relationship of This Work to Past Work

As stated in the Introduction, three important aspects of the Si/SiO_2 system require study: accurate thickness measurement as film thicknesses decrease, measurement of growth rates of very thin films, and the relationship between SiO_2 film thickness and optical properties. The following four sections of this dissertation address these topics. Each section is the manuscript of a paper to be submitted for publication in a refereed journal. The topics were selected to represent natural divisions of the project.

Section II is a physical description of the apparatus developed and built specifically for these measurements. The description is concluded with a discussion of the adaptation of this instrument for production line monitoring, both for process control feedback and *in situ* quality assurance. Further details of operation are included in the three Appendices following the last technical paper. These appendices are documentation for the operation of the instrument; they are somewhat informal to offset their technical complexity. Copies of the appendices also appear as .DOC files on the computer which controls the ellipsometer.

Section III presents Part I of a three-part scientific paper. The relationship between SiO_2 film thickness and refractive index is demonstrated. The data fit the model of Kalnitsky *et al.* (34) when a 1.5 nm thick transition region is assumed.

Section IV presents Part II, an x-ray photoelectron spectroscopy investigation of the thin film stoichiometry. This work grew out of a simple need to verify the "flash anneal" procedure crucial to this work, after Hopper *et al.* (20). The information gleaned from the chemical data provides corroboratory chemical and physical chemical evidence of the results shown in Part I, and allows us to offer a possible structural explanation for the observed optical behavior.

Section V presents Part III, which details a kinetic analysis of the initial stages of silicon oxidation. Two new oxidation regimes are reported which explain the rapid film growth unaccounted for by Deal and Grove. The abrupt change in reaction rate observed between these two stages is discussed in light of chemical data presented in Part II, and implications for device fabrication are explored.

References

1. R. S. Muller and T. I. Kamins. Device Electronics for Integrated Circuits, second edition. John Wiley & Sons, New York (1986), p. 57.

2. R. J. Archer, *J. Electrochem. Soc.*, **104**, 619 (1957).

3. R. J. Archer, *J. Opt. Soc. Am.*, **52**, 970 (1962).

4. A. M. Goodman and J. M. Breece, *J. Electrochem. Soc.*, **117**, 982 (1970).

5. E. Taft and L. Cordes, *ibid.*, **126**, 131, (1979).

6. D. E. Aspnes and J. B. Theeten, *Phys. Rev. Lett.*, **43**, 1046 (1979).

7. D. E. Aspnes and J. B. Theeten, *J. Electrochem. Soc.*, **127**, 1359 (1980).

8. J. B. Theeten and D. E. Aspnes, *Thin Solid Films*, **60**, 183 (1979).

9. G. A. Candela, D. Chandler-Horowitz, J. F. Marchiando, D. B. Novotnoy B. J. Belzer, and M. C. Croarkin, NIST Special Publication 260-109, U. S. Government Printing Office, Washington, DC (1988).

10. D. Chandler-Horowitz, *private communication.*

11. B. E. Deal and A. S. Grove, *J. Appl. Phys.*, **36**, 3770 (1965).

12. H. Z. Massoud, J. D. Plummer, and E. A. Irene, *J. Electrochem. Soc.*, **132**, 2685 (1985).

13. A. S. Grove, Physics and Technology of Semiconductor Devices. John Wiley & Sons, New York (1967), p. 32.

14. P. J. Jorgensen, *J. Chem. Phys.*, **37**, 874 (1962).

15. S. M. Hu, *Appl. Phys. Lett.*, **42**, 872 (1983).

16. M. Hamasaki, *Solid State Electron.*, **25**, 479 (1982).

17. A. G. Revesz and R. J. Evans, *J. Phys. Chem. Solids*, **30**, 551 (1969).

18. S. W. Ing, R. E. Morrison, and J. E. Sandor, *J. Electrochem. Soc.*, **109**, 221 (1962).

19. A. G. Revesz, B. J. Mrstik, H. L. Hughes, and D. McCarthy, *ibid.*, **133**, 586 (1986).

20. M. A. Hopper, R. A. Clarke, and L. Young, *ibid.*, **122**, 1216 (1975).

21. P. J. Jorgensen, *J. Chem. Phys.*, **49**, 1594 (1968).

22. D. O. Raleigh, *J. Electrochem. Soc.*, **113**, 782 (1966).

23. J. M. Gibson and D. W. Dong, *ibid.*, **127**, 2722 (1980).

24. P. G. Borden, *Appl. Phys. Lett.*, **36**, 829 (1980).

25. E. P. EerNisse, *ibid.*, **35**, 8 (1979).

26. H. Z. Massoud, J. D. Plummer, and E. A. Irene, *J. Electrochem. Soc.*, **132**, 1745 (1985).

27. A. Fargeix, G. Ghibaudo, and G. Kamarinos, *J. Appl. Phys.*, **54**, 2878 (1983).

28. C.-J. Han and C. R. Helms, *J. Electrochem. Soc.*, **134**, 1297 (1987).

29. C.-J. Han and C. R. Helms, *ibid.*, **135**, 1824 (1988).

30. A. G. Revesz, B. L. Mrstik, and H. L. Hughes, *ibid.*, **134**, 2911 (1987).

31. R. H. Doremus, *ibid.*, **134**, 2001 (1987).

32. E. A. Taft, *ibid.*, **134**, 475 (1987).

33. B. J. Mrstik, A. G. Revesz, M. Ancona, and H. L. Hughes, *ibid.*, **134**, 2020 (1987).

34. A. Kalnitsky, S. P. Tay, J. P. Ellul, S. Chongsawangvirod, J. W. Andrews, and E. A. Irene, *ibid.*, **137**, 234 (1990).

Section II

Experimental Aspects of the Growth of Thin SiO_2 Films

A paper to be submitted to *Review of Scientific Instruments*

A System for the Study of the Growth of Silicon Oxide Films with Real-Time Process Monitoring Capability

G. A. Danko

Department of Materials Science and Engineering

The Johns Hopkins University, Baltimore, Maryland 21218

ABSTRACT

An ultrahigh vacuum ellipsometric cell is described along with an ellipsometer capable of null or polarization modulation measurements. This system has been used to explore the initial stages of silicon oxidation at data acquisition rates of 10 points sec^{-1}. Data rates of 100 sec^{-1} are achievable. Hot stage design is discussed as it relates to specimen manipulation and accommodation of thermal expansion of the silicon substrate. Vacuum design, cleanliness and gas flow parameters are also discussed.

Introduction

Ellipsometry is an established technique for the characterization of surface oxides formed during silicon device processing (1, 2). However, within the past 13 years, a growing body of evidence has suggested that the silicon/oxide system cannot be explained by a simple substrate/film model (3, 4, 5). Our laboratory has sought to study the optical constants and kinetics of ultrathin (<10 nm) oxide films under actual conditions of growth. The development of a growth chamber/ellipsometer has been central to this work. We employed the technique of polarization modulation ellipsometry (6) for dynamic

measurements of *in situ* oxidation along with manual two-zone null ellipsometry to provide a check of accuracy and precision.

Ellipsometry is fundamentally a measure of two parameters, Δ and Ψ. Δ, the phase shift between the parallel and perpendicular components of a reflected light beam, is primarily an indicator of film thickness on the reflecting surface. The quantity Ψ is defined as the ratio of reflection coefficients $\tan^{-1}(|r_p|/|r_s|)$ of these same components, and is related to the refractive indexes of the substrate/film/medium system. These quantities are related to the reflectance ratio ρ by the equation (7)

$$\rho = r_p/r_s = \tan \Psi \exp[j\Delta].$$

These concepts are treated in excellent detail by Azzam and Bashara (8) and the reader is encouraged to seek this reference for a deeper understanding of optical theory. It is sufficient here to note that ρ is a function of many variables, $\rho = f(n_s, n_f, n_m, \lambda, \theta, d)$ where $n_i = N_i - jK_i$ represents the complex index of refraction of substrate, film, and ambient medium, respectively, λ is the wavelength of the probe light beam, θ is the angle of incidence between the light and specimen normal, and d is the thickness of the surface film.

Faced with the prospect of evaluating nine unknown quantities with only two pieces of information, the experimenter must take independent measure of several of them. We fix λ and θ, and note that $n_m = 1.0003 - 0.0j$ for air. Precise measurement of substrate temperature permits assignment of values for N_s and K_s from the data of van der Meulen and Hien (9). We have assumed that the oxide layer on silicon is transparent, i. e. has $n_f = N_f - 0.0j$, reducing our equation to the solution of two unknowns, d and N_f, with two data, Δ and Ψ. Solutions are sought numerically by iterating N_f in a computer model and

tracking the imaginary component of film thickness. Solutions are found when this value passes through zero since thickness is a physically real quantity.

Hopper, Clarke, and Young (10) have reported ellipsometric measurements of SiO_2 on hot silicon substrates; this project was inspired by their work. Hopper *et al.* relied on computer-controlled stepping motors to rotate the optical elements P (polarizer) and A (analyzer), continuously maintaining null intensity as recorded by a photomultiplier/amplifier. This electromechanical system required 10 seconds per data point, restricting their dynamic measurements to relatively slow chemical processes such as the parabolic growth regime of silicon oxidation, where thickness changes are on the order of 10^{-2} Ångstroms sec^{-1}.

The polarization modulation ellipsometer gives investigators the ability to record transient chemical phenomena. We have investigated the initial stages of SiO_2 film formation, identifying two new regions of oxide growth with sufficient time resolution to extract rate constants for the reactions (11, 12, 13). These data provide a more detailed understanding of the physical chemical behavior of the Si/SiO_2 system.

This paper describes the growth chamber, kinematic hot stage for specimen manipulation, and salient features of the ellipsometer. The system developed in our laboratory can be scaled up for process monitoring of silicon devices while the oxides are growing, providing feedback to the process controllers and ultimately increasing the production yields of complex solid-state devices.

1. Growth Chamber

A carefully controlled environment is paramount to the growth of device quality oxides. Cleanliness of the chamber and stage was assured by utilization of UHV-compatible

materials throughout. Most of these materials could withstand a standard 400° C bakeout, except for Viton® parts in the valves and some mica-glass parts on the hot stage. In consideration of these components, bakeout temperature was limited to 150° C. We adhered to standard UHV cleaning practice when dealing with the apparatus; specimens were cleaned according to electronic industry practice (11) and the oxidant was ultrahigh purity dry Matheson® oxygen.

The nature of our work required the capability of surface reduction as well as oxidation. Under the conditions of high temperature (1200° C) and low pressure (10^{-8} torr) active silicon oxidation is the thermodynamically favored reaction. Given the low partial pressure of oxygen in the chamber, the native air-formed SiO_2 film is reduced to SiO by the migration of silicon atoms from the substrate into the oxide. Since SiO can exist only as a gaseous phase, SiO sublimation and possible SiO_2 spalling from the specimen surface results. A fifteen second exposure to these conditions was adequate to achieve an essentially oxygen-free silicon surface, except for that small quantity of SiO which could reprecipitate on the surface during cooling.

The specimen chamber consists of a 304 stainless steel pipe spool 4" i. d. x 10.62" in length (see Figure 1). The long ends are 6" Conflat® rotatable flanges. Seven 2 3/4" Conflat flanges are arrayed about the chamber. Two flanges provide the laser entrance and exit ports, rotatable flanges are set at an angle of 70° and centered on a target point located 5.62" from the rear end of the chamber. One flange is directly above the sample target point; for the application described in this paper this accommodates a viewport, but can be used to access specimens in other experiments such as cleavage *in vacuo*. The other four flanges are arrayed on the top side of the chamber 30° off vertical, two behind and two in front of the target point. A piezoelectric leak valve and a spare flange are located behind the

target point. One flange ahead of target holds a thermistor vacuum gauge and a nude ion gauge. The other holds an up-to-air valve and a viewport used for pyrometry readings. Inside the chamber is a lug aligned with the rear face in the plane of the target point. This lug can be used as an aid in target point alignment or to provide mechanical stability for experiments such as cleavage *in vacuo*. The chamber is mounted on a 1/4-20 thread welded externally below the target point. A large ball-and-socket joint was fabricated to sit on the ellipsometer rail. The curvature of the joint is a 3 1/2" radius so that tilt of the entire chamber would be possible while maintaining reasonable specimen eucentricity.

The laser entrance and exit ports provide the beam path for the ellipsometer. The windows fabricated for these ports were among the most critical and expensive components of the system. The optical elements are 3/8" optically flat and annealed fused quartz, mounted on 1 1/2" of 1/8" thick fused quartz tubing. This tubing meets a graded glass seal to 7052 glass tubing which is, in turn, Kovar-mated to a 304 stainless steel bellows and Conflat flange. These windows are fully bakeable to 400° C. A brass frame surrounds the bellows to provide three point adjustment for window tilt. No measurable ellipticity was found in these windows, nor any measurable strain birefringence under vacuum.

The pumping system is shown in Figure 2. The chamber is mated to a 4" bellows which is hinged to limit movement to the vertical plane only. This permits movement of the vacuum system and some adjustment of vertical tilt of the growth chamber while preventing lateral tilt and collapse of the bellows (in the direction of the long axis) under atmospheric pressure differential. The chamber is cantilevered off of the pumping station which consists of a 4" gate valve and Santovac-filled expanded-mouth 4" oil diffusion pump with cold trap. This arrangement provides a pumping speed of 70 torr liter sec^{-1} at the specimen, with a chamber base pressure of 10^{-8} torr.

Temperature determinations were made with an optical pyrometer operated at a wavelength of 650 nm. We chose this method to avoid the uncertainties inherent in contact measurements, such as the appearance of cold spots where the thermocouple rests against the specimen, or insufficient heat transfer (due to poor mechanical contact) which could lead to underestimation of sample temperature. Thermocouple measurements also suffer from time lag due to the need for temperature equilibration of sample and thermocouple.

The physical layout of the chamber gives priority to the optical path of the ellipsometer and evacuation path of the pumping system; visual access to the sample face required the placement of a front surface mirror between the sample and a viewport window. Another mirror redirects the sample image to the pyrometer situated by the operator's station. This layout lets the operator control and monitor temperature while running the ellipsometer and vacuum system. It also necessitates correction of the pyrometer readings for absorption of light by the intervening optics. The non-blackbody (self luminous) nature of the experiment also requires corrections for the emissivity of the sample. A table of corrections was constructed accounting for both sources of error. Consequently, the low-temperature limit of our experiments, 800° C, was dictated by the low end of the pyrometer scale at 760° C.

II. Kinematic Hot Stage

Boron doped silicon substrates were clamped in compression between contact blocks and a DC potential applied. Specimen temperature was achieved by constant current ohmic heating. This simple experimental concept became a significant engineering challenge due to the following constraints: 1) the need for UHV compatibility and a low outgassing rate regardless of temperature, 2) ability to withstand high temperatures for long times, 3)

accommodation of specimen thermal expansion/contraction while maintaining the specimen in plane, 4) dimensional stability and low drift rates, 5) the ability to translate and tilt, 6) electrical isolation, and 7) minimal vibration.

The base of the stage is a 6" Conflat flange, which mounts onto the rotatable flange at the rear of the specimen chamber. This special order flange contains four Mini-Conflat flanges arrayed in a square. Three flanges accommodate bellows-sealed linear motion feedthroughs for specimen tilt and translation. The fourth holds a 9 pin instrument feedthrough. Six of the pins are used: two pins carry high current DC for specimen heating, two pins carry AC current for stage bakeout, and two pins connect to a chromel-alumel thermocouple mounted within the stage to monitor bakeout temperature.

A 304 stainless steel platen rests on the ends of the linear feedthroughs. The feedthroughs terminate in ball noses which engage the platen in a flat/cone/vee kinematic mount. Simultaneous movement of all three feedthroughs causes z-translation of the platen. With the upper left translator fixed, movement of the right translator only results in platen tilt about the vertical axis. Movement of the lower translator only results in tilt about the horizontal axis.

A bakeout heater is mounted on the back of the platen. It is a piece of nichrome wire wound into an element, mounted on Vycor standoffs, and insulated from the stage by mica sheets. A bakeout circuit uses a filament transformer to isolate the heater from line voltage in case of electrical leakage. The transformer is plugged into a lamp timer so that bakeout may be automatically performed at times when no one was in attendance. Temperatures of 150° C are attained in one hour of baking. At 10^{-8} torr, this was adequate for our purposes.

The front face of the platen carries the specimen mounts, as shown in Figure 3. One side of the specimen is mechanically fixed and electrically isolated while the other accommodates the thermal expansion and is electrically grounded. Two mica-glass (machinable ceramic) end blocks flank the platen to constrain the specimen loading train. The electrically isolated side of the loading train consists of an Inconel 600 contact held by two mica-glass knife edges. The positive lead from the specimen heat power supply is connected to this contact. The contact rests against one of the mica-glass end blocks. After repeated use, the Inconel gets pitted and contaminated with silicon; the contacts can be machined down and regrooved, and a shim inserted into the mica-glass end block to maintain loading spring pressure.

The other side of the loading train consists of a 304 stainless steel carrier slightly (0.002") wider than the Inconel contact. The contact is free to slide in the carrier but cannot slip out of plane. The thermal expansion coefficient of 304 is slightly larger than that of Inconel, hence as the temperature increases the contact will not bind in the carrier. A 304 stainless steel leaf spring (made from shim stock) resides between the contact and the mica-glass end block. The contact on this side of the specimen must be connected to the ground side of the power supply to minimize the chance of a high current ground loop through the instrument chassis.

There is a cutout in the platen behind the location of the specimen. As stated above, the pyrometric measurements rely on the assumption of non-blackbody conditions; this slot allows radiant energy to escape from the rear face of the sample without being reflected back through it.

III. Ellipsometer

The technique of polarization modulation ellipsometry (PME) was first described by Jasperson and Schnatterly (6) and has since been refined experimentally (14, 15) and mathematically (8, 16). PME is a radiometric technique which utilizes a series of lock-in amplifiers to determine the Stokes parameters N, S, and C (17) by deconvolution of a 50 kHz modulated light (laser) beam. In our instrument, the parameters N and S are measured by obtaining 2ω and ω signal components and ratioing them to standard values obtained through the use of calibration optics. The ellipsometric parameters Δ (phase shift) and Ψ (amplitude ratio) are derived from N and S by the relations (5)

$$\Psi = \frac{1}{2}\cdot\cos^{-1}(N)$$

$$\Delta = \sin^{-1}\left[\frac{S}{2\cdot\sin(\Psi)}\right]$$

These simple algebraic manipulations can be performed in real time by the data acquisition system. The technique falls prey to several instrumental limitations, however. Basic accuracy is limited by the digitization process: 1) the analog-to-digital converter employed has 12-bit resolution. This restricts our measurements to one part in 4096. Over the 360° range of a circle, basic accuracy is thus limited to 0.09°. This is almost an order of magnitude greater than that achievable in conventional null mode. 2) sequential scanning of the input channels can cause measurement errors in dynamic environments (e.g. a rapidly oxidizing surface). We have sidestepped this limitation by an electronic sleight-of-hand. Points acquired at intervals greater than 50 msec (20 sec^{-1}) utilize averaging of 96 samples. The sampling engine is set to 40 kHz which results in complete sampling in 2.4 msec. The lock-in amplifier time constants are set to 10 msec, thus any deflection of the

signals are so small as to be lost in the digitization noise. The 96 samples are loaded into computer memory as six channels summed 16 times to provide signal averaging. (Note the choice of 16 samples. This ensures that the digitized sum is not shifted left more than four bits. With 12 bit data, the total value of the sum does not exceed 16 bits, permitting the use of 16-bit unsigned integers without fear of fatal mathematical errors due to integer overflow.)

Our automated ellipsometer is a dual-mode instrument capable of high speed or high accuracy measurements. In the high speed mode, polarization modulation ellipsometry is employed. Data rates up to 1000 sec^{-1} are attainable in software, permitting study of transient surface phenomena such as gaseous oxidation reactions or double-layer formation in electrochemical systems. Of course, the comparatively long time constants imply a useful maximum data rate of 100 sec^{-1}; under real operating conditions, 10 sec^{-1} was adequate to follow the initial stages of thermal oxidation. We state the fastest number as a maximum for the current design.

Conventional null ellipsometry is used to acquire data of both superior accuracy and precision. Switchover between modes is rapid (though the operator must remember to do several things in the proper order), as a two-zone measurement (18) can be obtained in less than one minute.

The optical chain consists of a 5 mW He-Ne laser (λ = 632.8 nm), two beam steering mirrors, a linear polarizer (P), a Hinds PEM-80 photoelastic modulator (M), a gap in the rail for insertion of either a calibration polarizer and calibration $\lambda/4$ plate for PM mode or a $\lambda/4$ wave plate for null mode (Q), the specimen stage (S), two exit apertures, a linear analyzer (A), a periscope, and two detectors: a photodiode for PM mode and a photomultiplier for null mode.

The orientation of the optical elements for PME is straightforward. The polarizer P is oriented 90° to the plane of incidence. The modulator M is oriented with the long axis 45° to the plane of incidence. The modulated laser beam reflects from a sample at a 70° angle of incidence (this parameter is adjustable by the experimenter) through a linear analyzer A oriented to 315°. Insertion of a calibration polarizer CP between M and S allows collection of 2ω calibration data, and insertion of the λ/4 wave plate Q between M and CP allows collection of the ω calibration data.

Switchover to null measurement is achieved by turning off the modulator, inserting Q between M and S, flipping the periscope to divert the emergent beam to a photomultiplier, and adjusting P and A as required. When restarting M, a settling-in period of five seconds is all that is required before resuming PM operation.

The polarization modulation ellipsometer is driven by a free-running oscillator, the photoelastic modulator. The optical element resonates at approximately 50 kHz. The drive circuitry outputs two signals, f and 2f, which are used as phase references for the signal processing circuitry. Light detection and signal deconvolution are performed parasitically, in that nothing that the operator does to the signal detection chain has an effect on the ellipsometer itself.

The signal chain consists of a MRD555 photodiode, preamplifier with gain and offset corrections, a card rack containing the signal processing circuits, an AT&T 6300 personal computer, and diagnostic oscilloscope and voltmeter.

The photodiode is reverse-biased so that leakage current is proportional to incident light intensity. A 1 kΩ foot resistor limits the current and provides a voltage drop to ground that can be detected by an AD521KD instrumentation amplifier. Gain is switchable at x1 and

x10 and a 10-turn potentiometer permits voltage offset adjustment. The output is directed to the signal input on the card cage and to the oscilloscope.

The card cage houses ten circuit boards. Slot 1 contains a digital time base and digital distribution network designed for a modulator not used here, hence this card is not used. Slot 2 contains analog distribution amplifiers and the low pass filter which provides an average DC signal intensity used to normalize the frequency-derived signals described below.

Slots 3 and 5 contain the lock-in amplifiers (Evans Electronics model 4110) for the real and imaginary parts of the ω signal, respectively. They are fed intensity signals from the distribution amplifiers on card 2, and reference signals of frequency f from card 4, the ω phase control card (Evans Electronics model 4114). Amplifier gains and phase reference adjustments can be controlled from the front panel.

Slots 6 and 8 contain the lock-in amplifiers for the real and imaginary parts of the 2ω signal, respectively. They are fed intensity signals from the distribution amplifiers on card 2 and reference signals of frequency 2f from a small card attached directly to the card cage bus. The 2ω phase control card in slot 7 is not used because the input frequency limit on the phase control card is 50 kHz, whereas the 2f input is 100 kHz. The imaginary reference component is derived from the small circuit board which contains a 100kHz quadrature generator. The front panel phase controls are thus not active for the 2ω signals. This is not an important problem, for the vector magnitude of the signal is the only quantity needed.

Slot 9 contains the interface card to the computer. The low pass filtered DC and external input signals from card 2, real ω from card 3, imaginary ω from card 5, real 2ω from card

6, and imaginary 2ω from card 8 are transmitted to card 9 across the card cage bus. These signals are relayed down to the analog input multiplexer on the multifunction DASH-16 board (Metrabyte Corporation) in the AT&T 6300 personal computer. Also, front panel gain selection switch positions are sensed by card 9, multiplexed, and transmitted to the DASH-16 via digital I/O lines. This allows the computer to sense the amplifier gains and automatically scale the signal intensities.

Data acquisition routines were written using the C language. The program required approximately 40 pages of code and contains routines for calibration, manual acquisition, low- and high-speed automatic acquisition, data reduction, graphing, and file storage and retrieval. High-speed data acquisition (>20 samples sec^{-1}) requires direct memory access (DMA) data transfers from the acquisition card to computer memory. Several software interrupt drivers were written to achieve this throughput.

Scaling Up

The research-grade system described above holds the promise of process monitoring in commercial device fabrication facilities. We have successfully grown and analyzed oxide films *in situ*. From this experience, we propose two potential industrial applications for this instrument:

First, the null measurement mode permits product evaluation on the production line. This capability eliminates the need to pull wafers from the fabrication line for ellipsometric assessment. The possibility of *ex situ* contamination of the wafer is eliminated and production throughput is increased.

Second, polarization modulation mode can be employed to directly monitor oxide growth at critical steps of device processing. The data reduction algorithms can be embellished to

provide real-time oxide thickness measurements, requiring only temperature input to evaluate the data. Windows can be installed on critical process furnaces, and the ellipsometer beam reflected from a suitably prepared target point on the wafers within. The manual interventions required on the lab bench can easily be automated. Insertion of shutters and calibration optics can be relegated to servomotor control. Null measurements may also be attainable by servomotor control of the polarizing elements.

Acknowledgments

The author is indebted to Messrs. Mike Franckowiak and Walt Krug of the Department of Materials Science and Engineering for their talented machine shop work and suggestions for design refinements on the hot stage.

Many thanks to Dr. Shimson Gottesfeld of the Los Alamos National Laboratory and Dr. Deane Chandler-Horowitz of the National Institute of Standards and Technology for advice on the practical aspects of the application of polarization modulation ellipsometry.

This work was supported by the Office of Naval Research under grant number N00014-89-J-1265.

References

1. R. J. Archer, *J. Opt. Soc. Am.*, **52**, 970 (1962).

2. A. M. Goodman and J. M. Breece, *J. Electrochem. Soc.*, **117**, 982 (1970).

3. E. Taft and L. Cordes, *ibid.*, **126**, 131 (1979).

4. D. E. Aspnes and J. B. Theeten, *Phys. Rev. Lett.*, **43**, 1046 (1979).

5. D. E. Aspnes and J. B. Theeten, *J. Electrochem. Soc.*, **127**, 1359 (1980).

6. S. N. Jasperson and S. E. Schnatterly, *Rev. Sci. Instrum.*, **40**, 761 (1969).

7. L. Tronstad, *Z. Physik. Chem.*, **A142**, 241 (1929).

8. R. M. A. Azzam and N. M. Bashara, Ellipsometry and Polarized Light. North Holland, New York (1977).

9. Y. J. van der Meulen and N. C. Hien, *J. Opt. Soc. Am.*, **64**, 804 (1974).

10. M. A. Hopper, R. A. Clarke, and L. Young, *J. Electrochem. Soc.*, **122**, 1216 (1975).

11. G. A. Danko, J. Kruger, and A. G. Revesz, "Rapid Film Formation on Clean Silicon Surfaces I: Optical Properties", *to be submitted.*

12. G. A. Danko, J. Kruger, A. G. Revesz, and P. Searson, "Rapid Film Formation on Clean Silicon Surfaces II: Composition", *to be submitted.*

13. G. A. Danko, J. Kruger, and A. G. Revesz, "Rapid Film Formation on Clean Silicon Surfaces III: Kinetics", *to be submitted.*

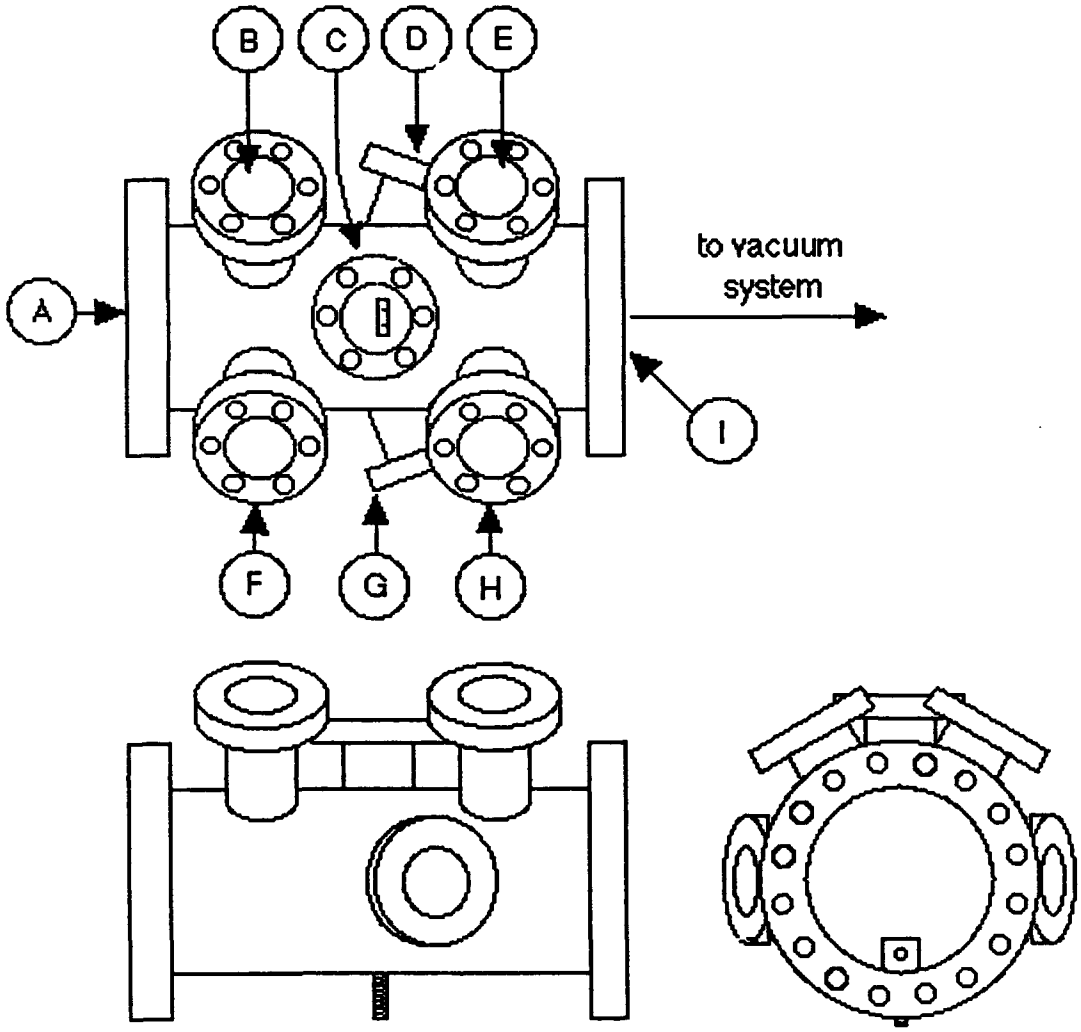
14. V. M. Bermudez and V. H. Ritz, *Appl. Opt.*, **17**, 542 (1978).

15. G. E. Jellison and F. A. Modine, *ibid.*, **29**, 959 (1990).

16. E. Huber, N. Baltzer, and M. von Allmen, *Rev. Sci. Instrum.*, **56**, 2222 (1985).

17. W. A. Shurcliffe, Polarized Light. Harvard University Press, Cambridge, MA (1962).

18. F. L. McCrackin, E. Passaglia, R. R. Stromberg, and H. L. Steinberg, *J. Res. NBS A*, **67A**, 363 (1963).



A - Hot Stage
B - Spare
C - Viewport
D - Laser Exit Window
E - Gauge Port
F - Piezoelectric Valve
G - Laser Entrance Window
H - Pyrometry Viewport
I - Pumping Port

Figure 1. Schematic drawing of the specimen chamber.

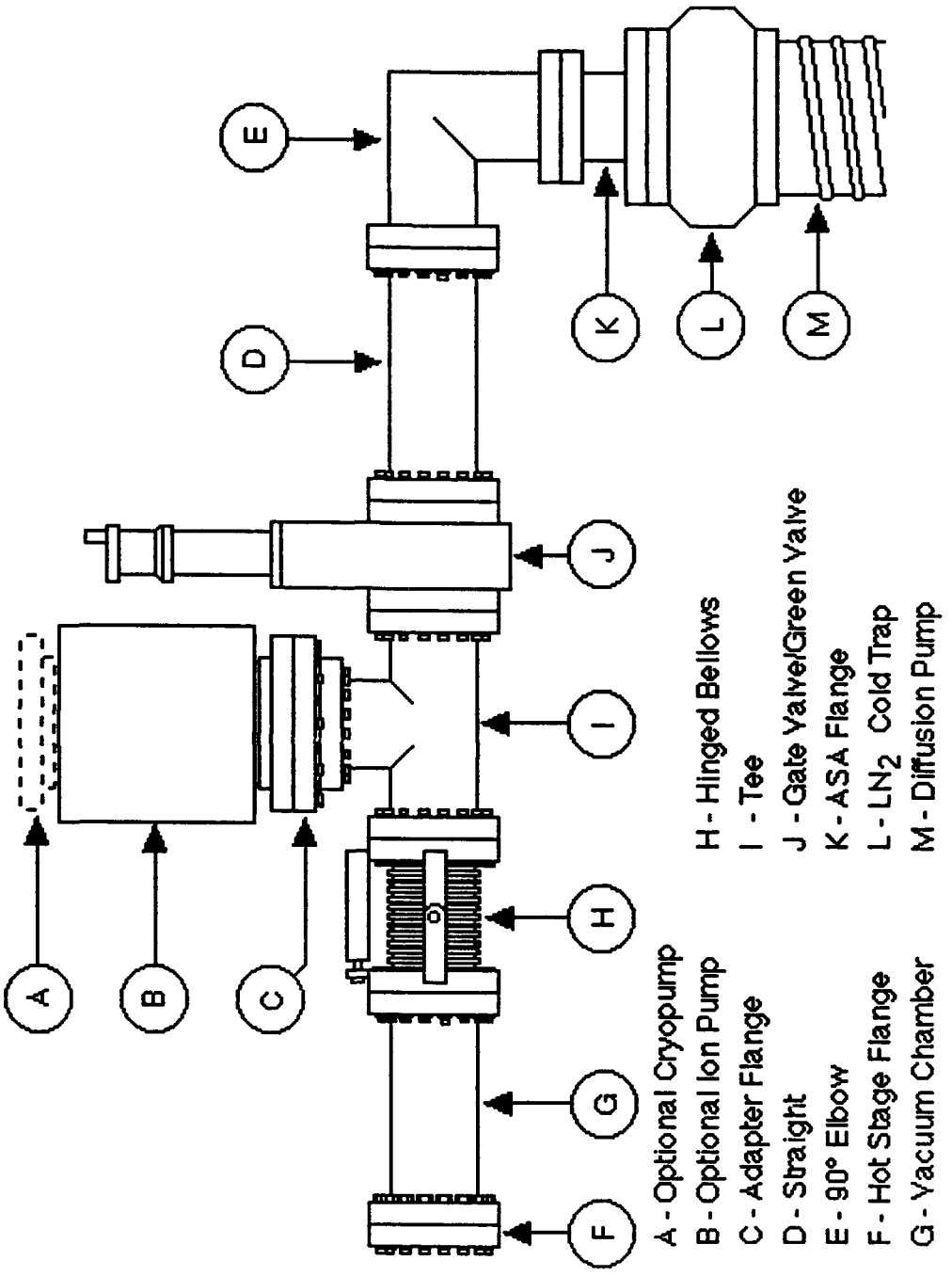


Figure 2. Schematic drawing of the vacuum system.

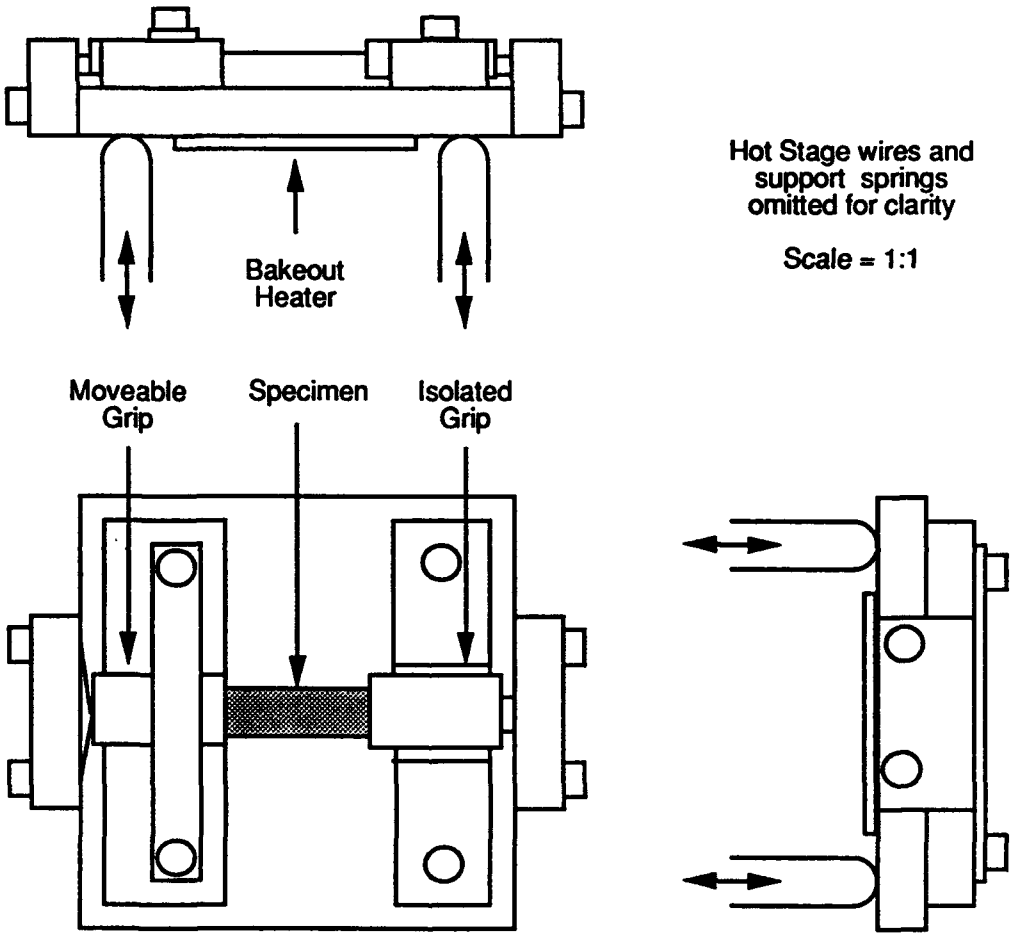


Figure 3. The kinematic hot stage.

Section III

Optical Properties of SiO_2 Films

A paper to be submitted to the *Journal of the Electrochemical Society*

Rapid Film Formation on Clean Silicon Surfaces
I: Optical Properties

G. A. Danko, J. Kruger

Department of Materials Science and Engineering

The Johns Hopkins University, Baltimore, Maryland 21218

A. G. Revesz

Revesz Associates, Bethesda, Maryland 20817

ABSTRACT

We present the results of an *in situ* ellipsometric growth study at $\lambda = 632.8$ nm with emphasis on the relationship of film refractive index to film thickness. A refractive index gradient was found that has good qualitative agreement with that reported by Kalnitsky *et al.* [*J. Electrochem. Soc.* **137**, 234 (1990)]. Refractive index behavior varies with process temperature but appears to be insensitive to substrate crystal orientation. Thin films of SiO_2 on silicon substrates often give large errors in ellipsometric measurement.

Introduction

Current MOS gate thicknesses are approaching 10 nm, and quantum tunneling devices will require SiO_2 thicknesses on the order of 3 nm. Lifetime predictions of radiation hardened devices are also becoming increasingly important. As data processing demands grow in

space applications, the complexity of on-board components must increase while retaining high reliability in the radiation environment of space. Moreover, continued evolution of silicon-based microelectronic circuitry will necessarily employ thinner SiO_2 masks for finer control of undercutting during the mask etching process. Efforts to assess the effects of point defects in protective SiO_2 are hampered, however, by insufficient data on the optical properties of very thin oxide films.

Early ellipsometric studies assumed a homogeneous SiO_2 film on silicon. Archer's work (1, 2) validated ellipsometry as a useful tool for semiconductor surface characterization. Zaininger and Revesz (3) cautioned against the validity of thick-film approximations used for ellipsometric analysis of thin film structures, but remained silent on the possibility of gradations in the optical properties near the film-substrate interface. Goodman and Breece (4) demonstrated reproducible thin oxide growth (~3 nm) on silicon at 600° C. Their time-thickness findings are sound, but no details of their ellipsometric data analysis was presented, leaving the present-day reader to question their modelling assumptions.

Taft and Cordes (5) performed an important etchback experiment in 1979 which led to the concept of an "interlayer" at the Si/SiO_2 interface. Their attempts to reconcile refractive index changes on a film measured after periodic HF etching succeeded by assuming an interlayer 0.6 nm in thickness with an index of refraction of $n = 2.8$. Confirmations of this finding were soon reported by others such as Aspnes and Theeten (6, 7). In ref. 7, the authors suggest that the transition region may be a graded structure rather than an abrupt interface, but this is based on a chemical mixing argument rather than direct experimental evidence. The interlayer model remains popular to this day, even though no physical structure has been identified to account for it.

We report the results of an experimental study of the optical properties of thermally grown oxides on silicon. A graded index of refraction near the Si-SiO_2 interface is demonstrated from high-temperature *in situ* ellipsometric measurements of the growing oxide film. This work is distinct from previous reports (5, 7, 8) whose authors inferred the interfacial behavior from room temperature etchback measurements of thick oxides as this work eliminates the chance of *ex situ* surface contamination.

This paper describes the behavior of the optical constants in thin SiO_2 films. Part II of this investigation details an XPS study of several thin (<10 nm) oxide films and will relate compositional variations to the variations in optical properties. Part III will discuss kinetics of the oxidation reaction from zero thickness through the first 10 nanometers of growth.

Experimental Procedures

The substrates used in this study were 2 inch diameter silicon wafers procured from Virginia Semiconductor, Inc. As received, the wafers were polished single side, boron doped (p-type) 0.1 ohm-cm, with an approximate thickness of 0.1 mm. Three orientations were used: Si(100), Si(111), and Si(110). All wafers were verified for orientation by transmission Laue measurements and wafers were randomly selected for four-point spreading resistivity checks.

20 mm x 5 mm samples were cleaved from the wafers, glued together (face-to-face and back-to-back) with Crystal Bond wax and edge polished with 300 grit SiC paper. The samples were washed in reagent grade acetone and individually cleaned by a modified version of the RCA standard clean (9): samples were boiled at 60° C for 15 minutes in H_2O:H_2O_2:NH_4OH 7:2:1, rinsed in deionized water, boiled for 15 minutes in

H_2O:H_2O_2:HCl 6:1:1, rinsed again and blown dry. Prior to use, each sample was to receive a dip in 10% HF. After gaining experience in initial experiments, the HF dip was discontinued, as the silicon repassivated before it could be transferred to the growth chamber. A recent report by Houssain *et al.* (10) validated this decision. Small batches (10 or less) of samples were prepared in this manner and stored until use under clean room conditions. In no case did the interval between batch cleaning and use exceed 10 days.

The growth chamber and radiometric polarization modulation ellipsometer (PME) are described in detail elsewhere (11). The light source is a 5 mW He-Ne (λ = 632.8 nm) laser. A photodiode detector was used for the radiometric measurements of PME mode. A photomultiplier tube was mounted perpendicular to the exit beam path; insertion of a mirror diverted the light beam to the photomultiplier so that conventional null ellipsometry could be performed. Switchover between modes was rapid, with about one minute required to interrupt PME operation, obtain the null measurement, and restart the PME. A 4 inch diameter Conflat®-sealed steel chamber was fabricated to sit on a modified Rudolph ellipsometer rail. Pressures of 1 x 10^{-8} torr were routinely achieved after overnight pumpdown and bakeout of the sample stage area. Samples were clamped between Pt-coated Inconel grips and heated by the application of an electric current, after the method of Hopper, Clarke, and Young (12). Once the samples were heated beyond thermal runaway (about 580° C) temperature could be reliably controlled by adjusting the current limit on the power supply. Temperatures up to the melting point of silicon were achievable, and accuracy and stability were within our ability to resolve with an optical pyrometer.

Optical pyrometry (650 nm wavelength) was employed to determine specimen temperature. Monitoring a self-luminous body *in vacuo* requires careful attention to two correction factors. First, direct observation of a sample in a vacuum system is not possible. At least

one window must be present to maintain vacuum, and attenuation of the light by that window will result in an underestimation of sample temperature. The construction of our vacuum chamber required the use of one window and two mirrors which further increased the potential for measurement error. Experimental attenuation measurements and Wien law calculations (13) were performed to create a table of temperature corrections ("apparent" temperatures) to account for the light loss. Second, emissivity corrections are required for the non-blackbody conditions in our cell. Both the real and imaginary parts of the complex index of refraction of silicon $n_s = n_s - jk_s$ (the subscript 's' refers to the silicon substrate) are functions of temperature; van der Meulen and Hien (14) published useful values for n_s and k_s of silicon at elevated temperatures. A non-linear least squares fit to their data was incorporated into a small computer program which could provide interpolated n_s and k_s for given apparent temperatures. Emissivity values, ε, were calculated from these optical constants by the equation of Sato (15):

$$\varepsilon = \frac{4n}{(n+1)^2+k^2} \qquad [1]$$

Wien law calculations based on this emissivity produced a "true" surface temperature, which was then fed back into the program to refine selection of the optical parameters. Convergence to true surface temperature occurred after three program iterations.

Sato demonstrated that the spectral emissivity varies only slightly in the 650 nm region, thus we felt justified in applying the same optical constants for temperature determination at 650 nm and ellipsometry at 632.8 nm. Computed parameters for temperature, n_s, k_s, κ_s ($\kappa_s = \frac{k_s}{n_s}$), and ε are summarized in Table I.

The ultrahigh vacuum conditions attainable in our growth chamber were required to carry out the final specimen cleaning. Specimens underwent a flash anneal at 1200° C for 15

seconds at a pressure of 10^{-8} torr. At this temperature silicon monoxide becomes the thermodynamically favored species (16) due to silicon diffusion into the oxide film. The thermodynamically favored form of the monoxide is gaseous; the production of a bare substrate surface occurs through the process of sublimation. A flash anneal performed *in situ* in an x-ray photoelectron spectrometer confirmed that oxide removal was essentially complete with only a fraction of a monolayer remaining. Any remaining oxygen is probably due to reprecipitation of SiO on sample cooling. At pressures of 10^{-8} torr the sample surface can be expected to take up a full monolayer of oxide in approximately 100 seconds. This time interval was sufficient to obtain a manual null ellipsometer measurement, establish a process temperature, start rapid ellipsometer data acquisition, and introduce ultrahigh purity Matheson oxygen into the growth chamber.

Oxidation of three silicon surface orientations was investigated at five temperatures from 800° C to 1000° C. All experiments were performed at 1 atm pressure in a dry oxygen ambient. Oxidation times varied from 20 minutes to over three days (one ellipsometric period). A total of 58 experiments were performed.

Results and Discussion

The data presented here were reduced from two-zone null ellipsometry. The two-zone configuration eliminates systematic errors in the optical chain and permits for mathematical correction of specimen tilt (17); small tilt errors invariably occur on sample heating due to thermal expansion of the specimen and resultant "settling in" of the grips. Only the results of the null ellipsometry will be discussed in Part I of this series.

Two parameters are obtained during an ellipsometric measurement. One is Δ, the phase shift of the components of the probe beam that lies parallel (r_p) and perpendicular (r_s) to the plane of incidence, defined by the expression $\Delta = (\beta_r - \beta_s)_{reflected} - (\beta_r - \beta_s)_{incident}$. The other datum is Ψ, the amplitude ratio $\Psi = \frac{|r_p|}{|r_s|}$ of said components. These variables provide a complete description of the reflectance ratio, ρ, of a surface by the relation $\rho = \tan\Psi \exp[j\Delta]$.

Experimental Δ, Ψ data are shown in Figures 1a-e. The circles represent theoretical Δ, Ψ pairs based on the assumption of a nonabsorbing film of refractive index $n_f = 1.461$ for all temperatures. Substrate refractive indexes were those given in Table I. The data for Figures 1a-d depart significantly from theory for thin films but approach the theoretical predictions at higher thicknesses. Notably, the data of Figure 1e lie much closer to theory. These experiments were performed at 1000° C, well above the viscous flow threshold temperature for SiO_2 (18, 19). We infer that the lower temperature samples may in some way be constrained by interfacial energy. Only as the film grows to several tens of nanometers thickness does the SiO_2 volume energy overtake the interfacial surface energy and present an oxide with the expected optical behavior.

The Δ, Ψ data of Figure 1 were input to McCrackin's ellipsometry program (20) via a routine that fits both film thickness, d, and film refractive index, n_f. The imaginary component of film thickness is monitored; since thickness is a real quantity, zero imaginary thickness is the program's requirement for a possible solution. A change in sign of the imaginary film thickness yields an approximate root whose accuracy is dependent on the fineness of the search grid.

Results of these calculations are presented in Figure 2. Each of these measurements represent the optical behavior of the entire film thickness at the time of measurement. No multilayer models were generated to interpret the behavior. Selected calculations using another routine were run to verify the fit of the model. The power of decision was taken away from the computer and solutions were printed for every value of n_f input. The solutions presented in Figure 2 are the best fits to our experimental data; imaginary thickness typically passes within 0.02 Ångstroms of zero and the magnitudes of $\delta\Delta$, $\delta\Psi$ ($\Delta_{theory} - \Delta_{experiment}$, $\Psi_{theory} - \Psi_{experiment}$) are always less than 0.002°.

The latter computations also yield a surprising result: a second "solution" appears to exist at $n_f = 2.8 \pm 0.02$. This refractive index value appears in all computations regardless of film thickness. Large errors in $\delta\Psi$ and imaginary thickness are observed with only $\delta\Delta$ passing through zero. The $\delta\Psi$ errors are often 1° or larger and bear no correlation to either real or imaginary film thickness. This finding leads us to conclude that the interlayer film model so commonly used is best explained as an experimental artifact. Archer (1) may have unknowingly alluded to this when he reported variations in Ψ measurements (which are strongly influenced by n_f) whereas his fits to Δ at multiple angles of incidence were excellent.

Recent work by Kalnitsky *et al.* (8) proposes a graded refractive index described by the equation

$$n(x) = 2.44 \exp[-0.5 \cdot (X/S)^2] + 1.46 \qquad [2$$

where X is distance measured from the interface, and S is the process related parameter describing the transition region width. Our results agree with their predictions in the thin film regime (<10 nm) where the refractive index is changing rapidly. A reasonable fit is

achieved for S = 5, which Kalnitsky *et al.* state is equivalent to a transition region width of 1.5 nm. The model does not adequately describe the turn in the experimental data between 100 Å and 400 Å.

Conclusions

1. The refractive index of thermally grown SiO_2 thin films on silicon substrates is non-linear with increasing film thickness. A refractive index gradient is evident which can be modelled by an empirical equation. The interface transition region appears to be about 1.5 nm thick.

2. The interlayer two-film model does not appear valid. We propose that it is an artifact of the data analysis.

3. The optical behavior of the oxide is not dependent on substrate orientation.

4. The optical behavior of the oxide may be dependent on process temperature.

In Part II of this paper we will present data from an x-ray photoelectron spectroscopy study and relate the chemical information to the thickness and structure of the transition region.

Acknowledgments

This work was supported by the Office of Naval Research under grant no. N00014-89-J-1265.

References

1. R. J. Archer, *J. Electrochem. Soc.*, **104**, 619 (1957).

2. R. J. Archer, *J. Opt. Soc. Am.*, **52**, 970 (1962).

3. A. M. Goodman and J. M. Breece, *J. Electrochem. Soc.*, **117**, 982 (1970).

4. K. H. Zaininger and A. G. Revesz, *RCA Rev.*, **25**, 85 (1964).

5. E. Taft and L. Cordes, *J. Electrochem. Soc.*, **126**, 131 (1979).

6. D. E. Aspnes and J. B. Theeten, *Phys. Rev. Lett.*, **43**, 1046 (1979).

7. D. E. Aspnes and J. B. Theeten, *J. Electrochem. Soc.*, **127**, 1359 (1980).

8. A. Kalnitsky, S. P. Tay, J. P. Ellul, S. Chongsawangvirod, J. W. Andrews, and E. A. Irene, *ibid.*, **137**, 234 (1990).

9. W. Kern and D. A. Puotinen, *RCA Rev.*, **31**, 187 (1970).

10. S. D. Houssain, C. G. Pantano, and J. Ruzyllo, *J. Electrochem. Soc.*, **137**, 3287 (1990).

11. G. A. Danko, "A System for the Study of the Growth of Silicon Oxide Films with Real-Time Process Monitoring Capability", *to be submitted.*

12. M. A. Hopper, R. A. Clarke, and L. Young, *J. Electrochem. Soc.*, **122**, 1216 (1975).

13. W. P. Wood and J. M. Cork, Pyrometry. McGraw-Hill, New York (1927).

14. Y. J. van der Meulen and N. C. Hien, *J. Opt. Soc. Am.*, **64**, 804 (1974).

15. T. Sato, *Japan. J. Appl. Phys.*, **6**, 339 (1967).

16. E. A. Gulbransen and S. A. Jansson, *Oxid. Met.*, **4**, 181 (1972).

17. F. L. McCrackin, E. Passaglia, R. R. Stromberg, and H. L. Steinberg, *J. Res. NBS A*, **67A**, 363 (1963).

18. E. A. Irene, E. Tierny, and J. Angilello, *J. Electrochem. Soc.*, **129**, 2594 (1982).

19. E. Kobeda and E. A. Irene, *J. Vac. Sci. Technol. B*, **6**, 574 (1988).

20. F. L. McCrackin, NBS Technical Note 479, U. S. Government Printing Office, Washington, DC (1969).

$T_{pyrometer}$	$T_{apparent}$	T_{true}	n_s	k_s	κ_s	ε
		RT	3.875	0.018	0.0046	
761	775	800	4.206	0.124	0.0297	0.620
808	824	850	4.234	0.139	0.0327	0.618
853	870	900	4.262	0.153	0.0360	0.615
899	917	950	4.291	0.170	0.0395	0.612
945	965	1000	4.321	0.188	0.0435	0.610

Table I. Optical parameters used in this work. See text for details.

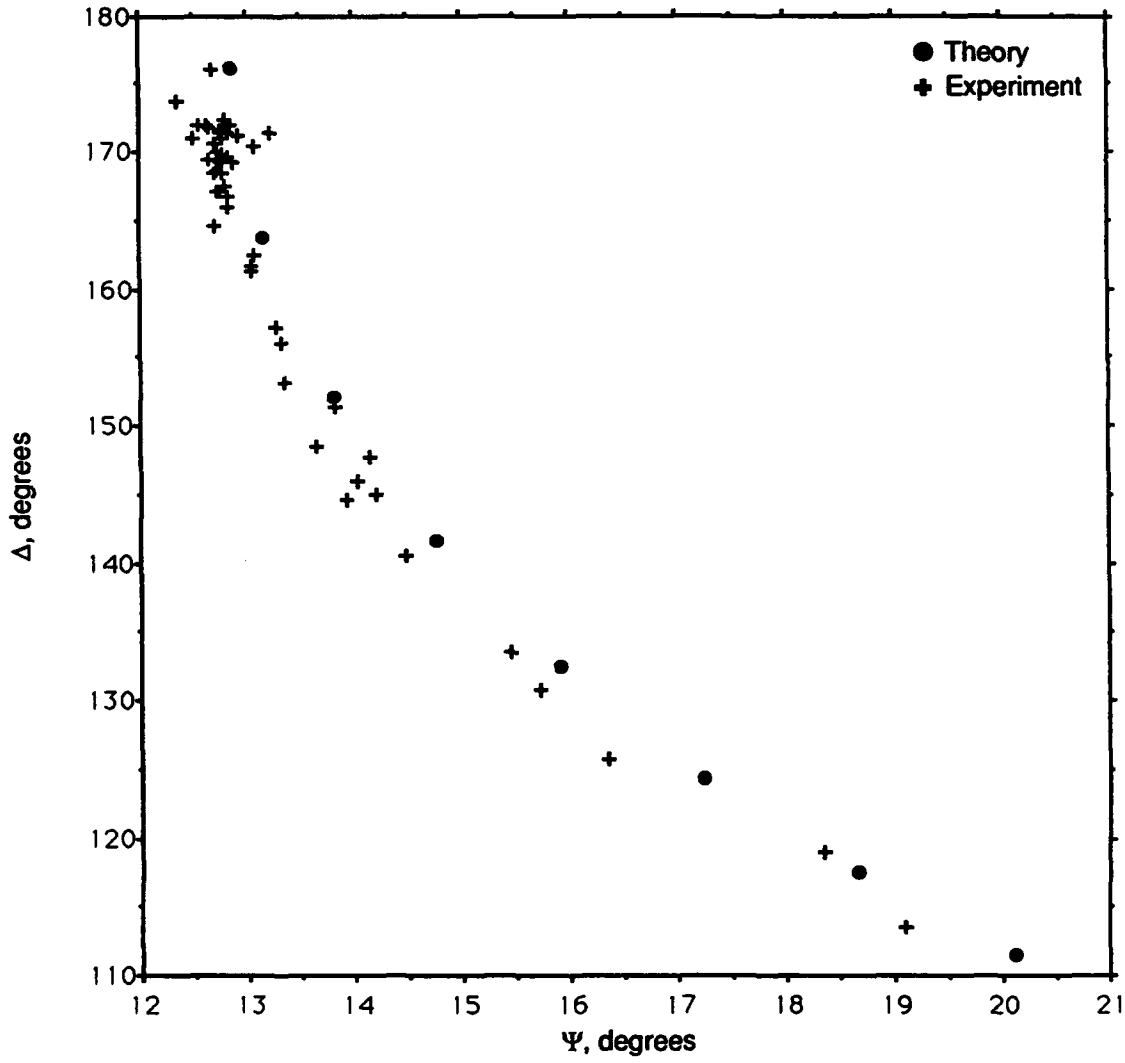


Figure 1. Δ vs. Ψ maps of ellipsometric data. (a) Oxidation temperature: 800° C. Points represent combined data from 20 experiments.

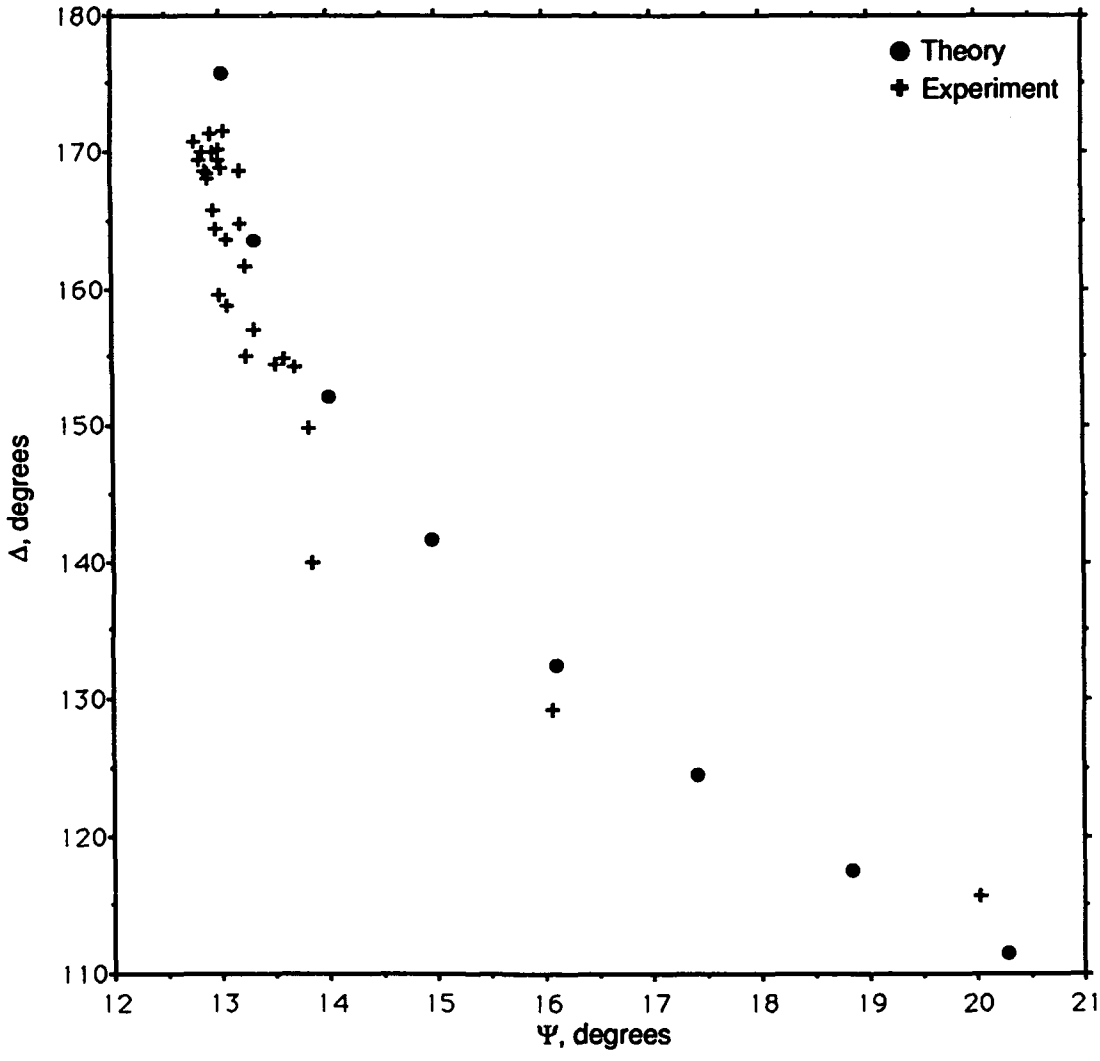


Figure 1. Δ vs. Ψ maps of ellipsometric data. (b) Oxidation temperature: 850° C. Points represent combined data from 14 experiments.

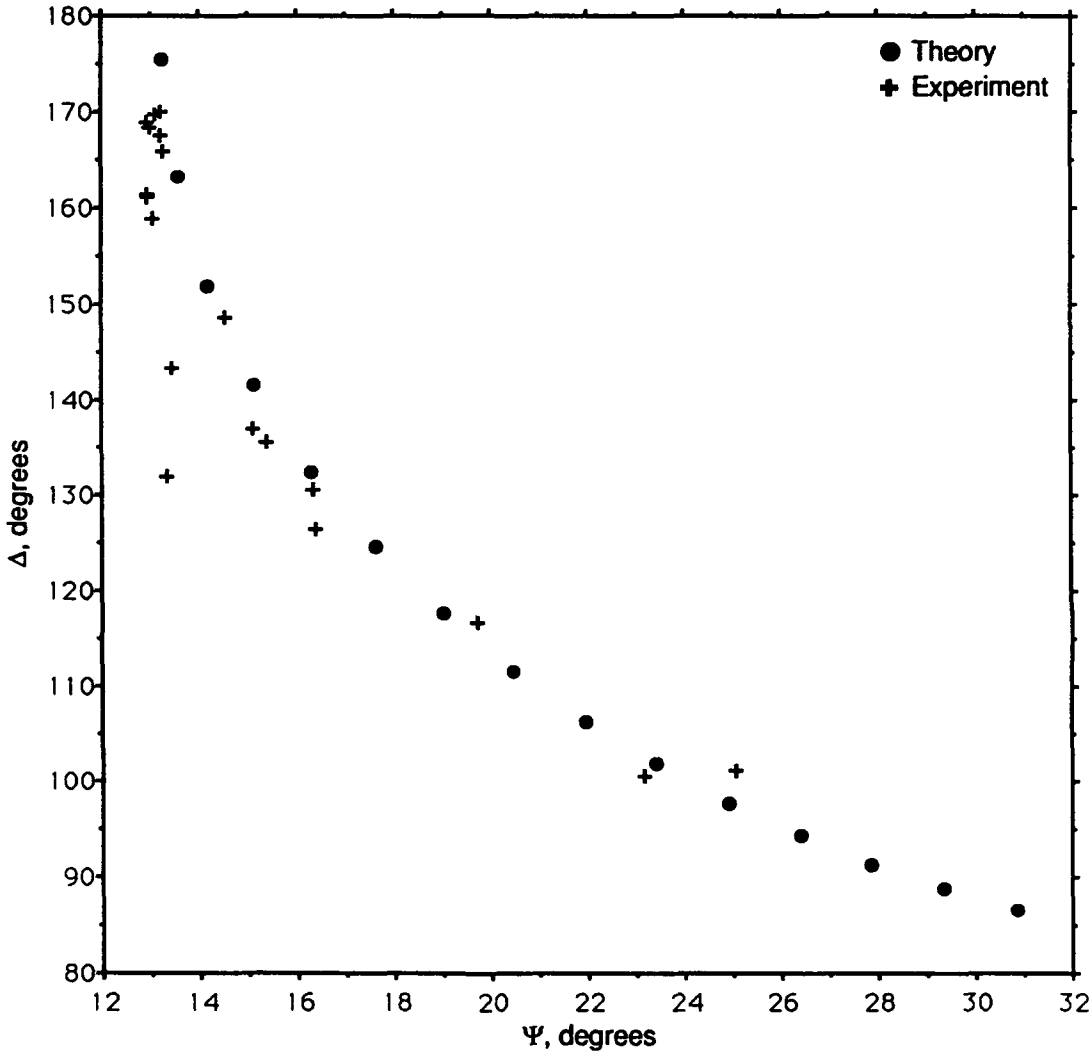


Figure 1. Δ vs. Ψ maps of ellipsometric data. (c) Oxidation temperature: 900° C. Points represent combined data from 7 experiments.

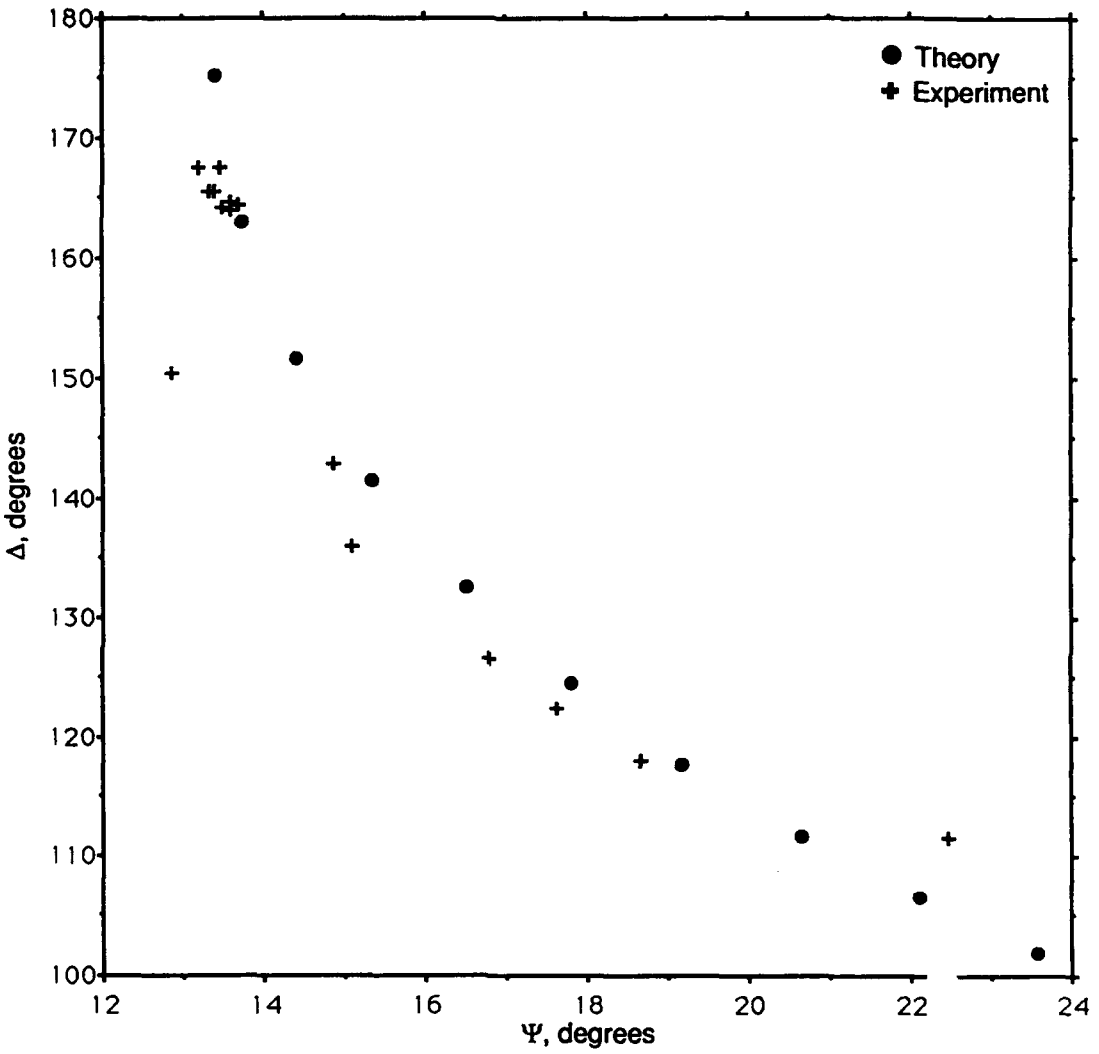


Figure 1. Δ vs. Ψ maps of ellipsometric data. (d) Oxidation temperature: 950° C. Points represent combined data from 8 experiments.

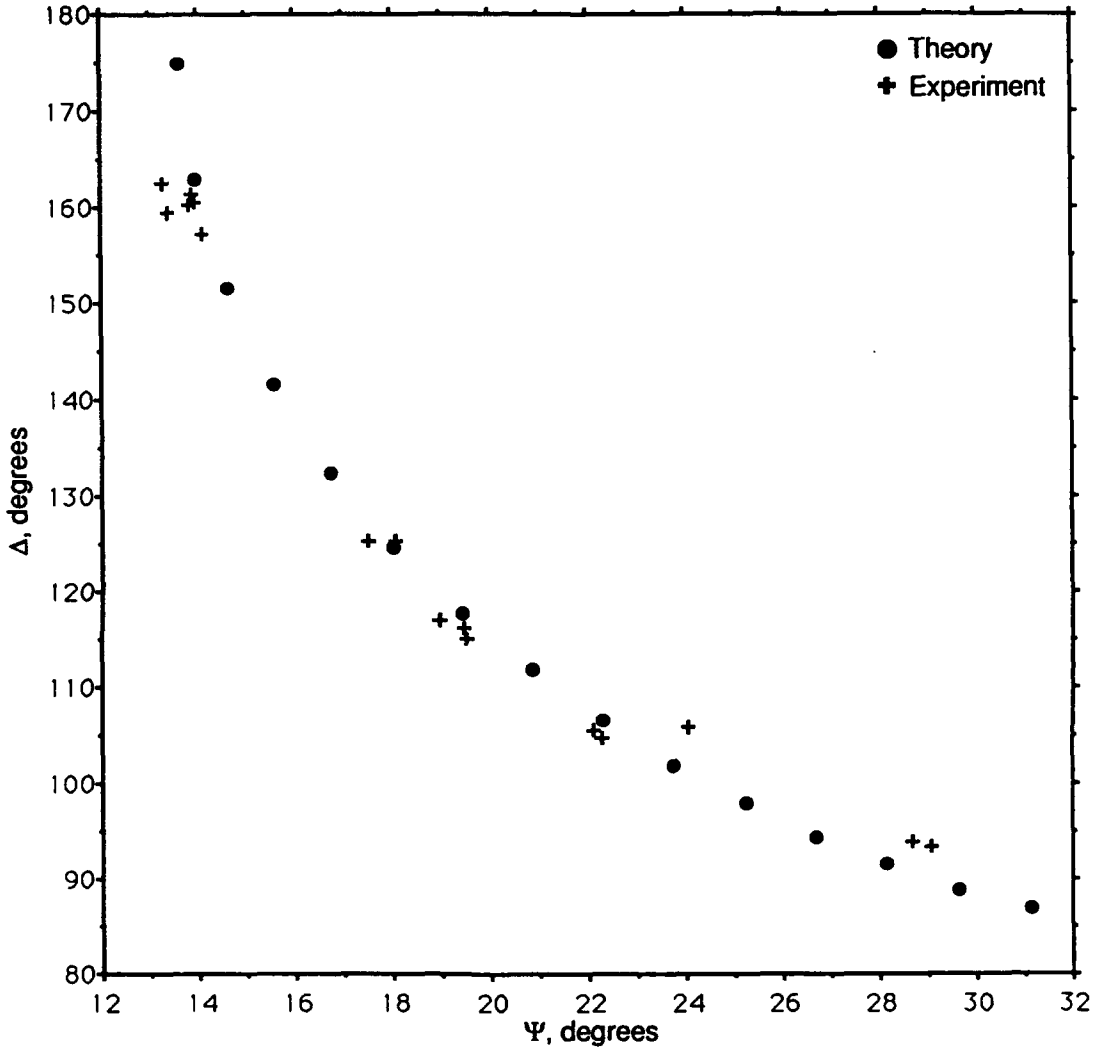


Figure 1. Δ vs. Ψ maps of ellipsometric data. (e) Oxidation temperature: 1000° C. Points represent combined data from 6 experiments.

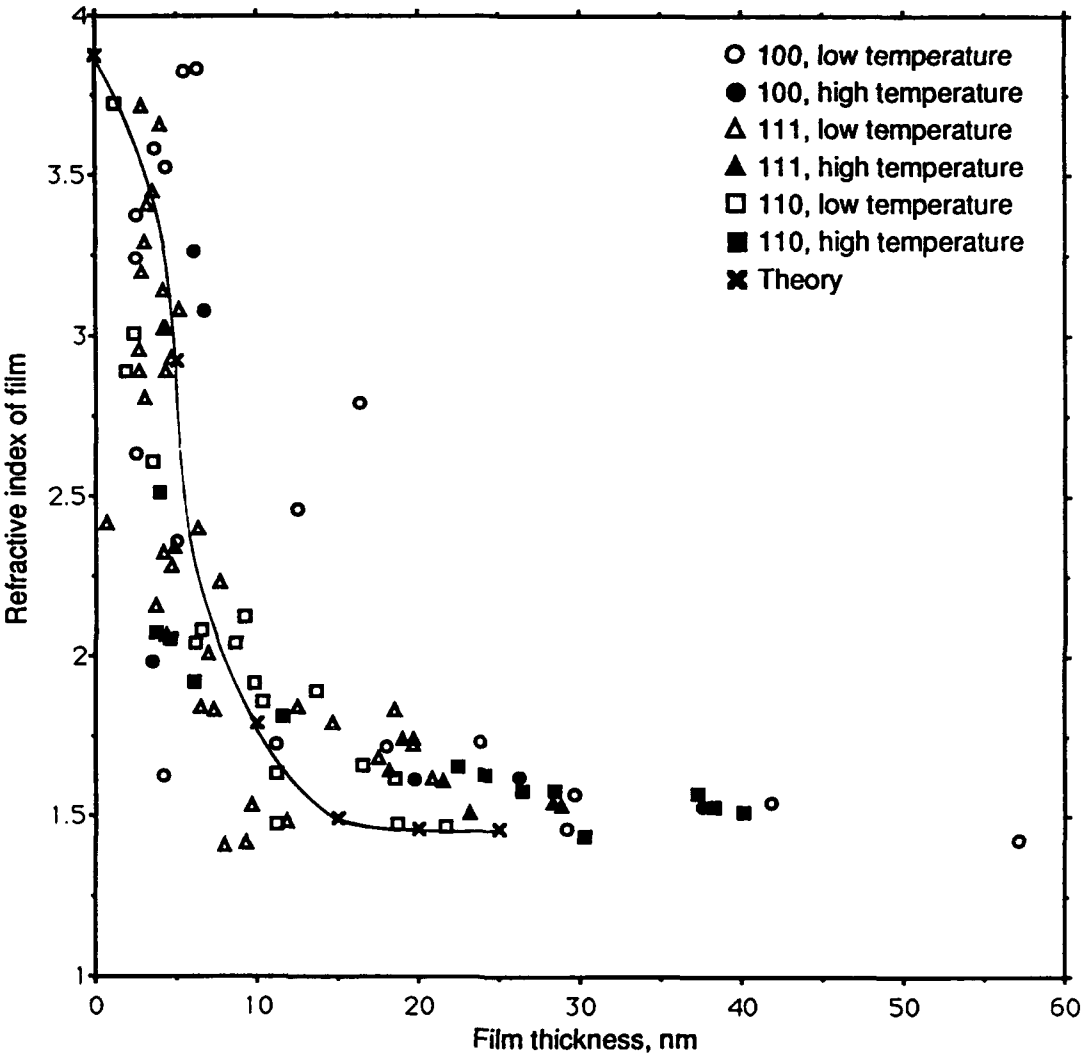


Figure 2. SiO_2 refractive index vs. film thickness for various temperatures and Si substrate orientations. The solid points represent data from experiments run at or above the temperature required for viscous flow of the oxide. Predictions from equation [2] are overlaid for comparison.

Section IV

Composition of Thin SiO_2 Films

A paper to be submitted to the *Journal of the Electrochemical Society*

Rapid Silicon Dioxide Film Formation on Clean Silicon Surfaces II: Composition

G. A. Danko, J. Kruger

Department of Materials Science and Engineering

The Johns Hopkins University, Baltimore, Maryland 21218

A. G. Revesz

Revesz Associates, Bethesda, Maryland 20817

P. Searson

Department of Materials Science and Engineering

The Johns Hopkins University, Baltimore, Maryland 21218

ABSTRACT

Ultrathin (<10 nm) oxide films were thermally grown on Si(111) substrates. Angle-resolved XPS measurements reveal a coesite-like behavior near the silicon/oxide interface. A persistent high oxygen/silicon ratio exists irrespective of oxide thickness or postoxidation annealing *in vacuo*. The results are employed to explain observed variations in the refractive index of thin SiO_2 films.

Introduction

The aim of this study was to gain a more detailed picture of the composition and structure of the SiO_2 films studied in Part I and to relate this information to the optical constant measurements reported there. Silicon (111) samples were oxidized using growth techniques described elsewhere (1). Eight specimens were oxidized under various pre- and postgrowth conditions: first, a comparison of oxide composition between flash annealed precursors and intact native oxide precursors was proposed, and second, with regard to the graded refractive index behavior discussed in Part I, several of the thermally oxidized specimens were annealed *in vacuo* to test the viscoelastic stress-relief hypothesis of Irene *et al.* (2). These samples were in turn analyzed by angle-resolved x-ray photoelectron spectroscopy.

Experimental Procedures

Nine specimens were cleaved from a 2-inch diameter silicon (111) wafer procured from Virginia Semiconductor, Inc. The wafer, as with others used throughout this study, was B-doped 0.1 ohm-cm and polished on one side only. 5 x 20 mm specimens were cleaved, edge polished, and RCA cleaned in the manner described previously (1). Fresh solutions were prepared for each sample to ensure cleanliness. Six of the samples were batch stored under clean room conditions, while the remaining three were cleaned immediately (<5 minutes) before use. Samples were individually oxidized in an ultrahigh vacuum growth chamber/ellipsometer, then transferred to a Perkin-Elmer 5100 x-ray photoelectron spectrometer. During transfer, the samples were exposed to room air for periods of up to 15 minutes. Surface contamination cannot be ignored in analyzing the data; in fact, carbon

deposition from pump oils was used to correct for elemental peak shifts from specimen to specimen.

Three different experimental protocols were followed. The first was a series of control experiments which repeated the steps followed in the optical constants/kinetics work reported in Parts I and II of this paper. We shall refer to these experiments by the acronym *Flox*, as explained in more detail below. The other protocols, *Noflox* and *Anneal*, were constructed to investigate pre- and post-growth effects on the behavior of the oxide layers.

The acronym *Flox* stands for *Fl*ash annealed and *ox*idized. The flash anneal effects removal of the native surface oxide, allowing our oxidation experiments to start from zero thickness. Under suitable conditions of temperature and pressure, SiO becomes the thermodynamically favored species (3). A 15 second exposure to a temperature of 1200° C and pressure of 10^{-8} torr was chosen for this work. SiO is gaseous under these conditions, subliming from the silicon substrate and possibly spalling off overlying layers of SiO_2. The SiO was pumped away although much of it condensed on the chamber walls, providing a gettering action to further lower the chamber partial pressures of O_2 and H_2O. A temperature of 850° C was established and 1 atm of ultrahigh purity oxygen was introduced. Specimens were oxidized to attain a range of thicknesses below 10 nm. Two-zone null ellipsometric measurements were performed between the different stages of surface modification. Polarization modulation ellipsometry was employed to follow oxidation *in situ*.

The *Noflox* protocol is similar to the first, but with *No fl*ash anneal, followed by thermal *ox*idation. For these experiments, specimens were individually subjected to a modified RCA clean immediately prior to insertion in the growth chamber. The native oxide was

subject to a 150° C bakeout for one hour at 10^{-8} torr. Oxidation conditions and post-growth specimen handling were similar to the first protocol.

The *Anneal* protocol involved flash anneal and oxidation as in the first protocol. Oxidation was halted by cooling the samples to room temperature. The growth chamber was then evacuated of oxygen to 10^{-2} torr and flushed three times with dry nitrogen. The chamber was evacuated into the 10^{-8} torr range and the sample subjected to a 15 second anneal at 1100° C. Postanneal specimen handling was the same as in the other protocols.

Samples were examined in the XPS while still "fresh", i. e. within one half hour of oxide formation. The Perkin-Elmer 5100 spectrometer was operated at 0.5 eV resolution at full width half-maximum (FWHM) at a pressure of less than $5x10^{-9}$ torr. Incident x-radiation was generated by a MgKα source (1253.6 eV) operated at 300 W. The spherical capacitor energy analyzer was operated with a constant pass energy of 35.75 volts. Five takeoff angles, θ = 15°, 30°, 45°, 60°, and 75° were utilized to provide depth-sensitive information. This angle is defined as that between the sample surface and the axis of the electron optics. Survey scans were performed on each sample to determine elemental composition; samples were then analyzed at 0.5 eV resolution on all observed elemental peaks. An *in situ* flash anneal and room temperature oxidation experiment was carried out to verify the completeness of surface reduction. Electronic grade quartz was also analyzed to check the accuracy of the instrument. The sample survey scans were from zero to 1000 eV to reveal elemental composition. The Si $2p_{1/2}$, O1s, and C1s edges were then examined with high resolution (0.5 eV/channel) windows. Nickel was detected on one sample, and the $Ni2p_{1/2}$, $Ni2p_{3/2}$ lines were recorded.

An *in situ* flash anneal was performed to verify the completeness of surface reduction. The hot stage platen (4) was removed from its kinematic supports in the growth chamber and

fitted to the XPS sample stage, which required breaking XPS vacuum. The sample was freshly cleaned upon insertion, but had to endure a 400° C spectrometer bakeout. Angle-resolved XPS measurements were made on the native oxide. The specimen underwent a 15 second flash anneal, was cooled to room temperature, and subjected to an identical series of measurements. Ultrahigh purity oxygen was admitted to the XPS chamber for 15 minutes, resulting in a room temperature oxidation reaction on the bare surface. Angle-resolved XPS measurements were performed subsequent to pumpdown and another 400° C system bakeout.

Spectrometer benchmarks were established by analyzing an electronic-grade quartz resonator. This specimen was immersed for one minute in a 10% HF dip, rinsed with 10^{18} ohm-cm water, blown dry and inserted into the XPS. This procedure provided us with data from a reliable SiO_2 standard that had been subjected to the same room-air contamination as the oxidized silicon substrates.

Results and Discussion

One of the important aspects of this study is that it is aimed at preventing surface modification during analysis. As discussed by Grunthaner *et al.* (5), depth profiling techniques (Argon ion milling or wet chemical etching) can seriously alter the structure and composition of the near-surface layers of a specimen. In the case of ion milling, ionization damage and momentum-induced or knock-on damage can produce changes in both structure and specimen stoichiometry, as well as changes in space-charged structures such as stacking faults.

Wet techniques invite the introduction of foreign contaminants, notably H_2O and protonated species that may radically affect the specimen surface chemistry. Wet techniques also suffer from inaccuracies in etch rate. Such errors arise from improperly mixed etchants, inaccurate timing of reaction, improper solution agitation, and so forth, all of which can affect solution activity on a microscopic scale. Even small variations in etch rate (e. g. 1 nm sec^{-1}) would be disastrous on a film less than 10 nm in thickness.

Data from the *in situ* flash anneal are shown in Figure 1. The disappearance of the Si^{+4} (we refer here to formal oxidation states, not the total transfer of charge, which is somewhat less due to the partially covalent nature of the Si-O bond) peak in Figure 1a is clearly indicative of the removal of surface oxide. The nearly complete removal of oxygen shown in Figure 1b indicates removal of the majority of the oxide, perhaps with reprecipitation of SiO on cooling.

Nickel was detected on this sample after the flash as shown in Figure 2. The Ni_{2p} doublet has a peak separation measured at 17.5 eV, indicative of metallic nickel. We suspect that small beads of molten nickel were ejected by a small arc between the silicon wafer and the Inconel grips, solidifying as spheres or islands on the silicon surface. The integrated Ni_{2p} signal amounted to less that 0.1% of the total integrated elemental signal and did not change with takeoff angle, indicative of a spherical particle geometry. Based on this information, we can state that 1) nickel contamination should not measurably affect the optical measurements, 2) the nickel distribution is unlikely to either catalytically enhance or poison SiO_2 formation.

On readmission of O_2, some native oxide formation did occur. Figure 3a reveals mixed silicon valences similar to those reported by Borman *et al.* (6); figure 3b shows a split

oxygen peak that will be discussed in detail below. Figure 3c demonstrates the invariance of the Ni_{2p} doublet versus the takeoff angle. The fact that the signal is still metallic after oxygen admission strongly suggests that one or a few relatively large beads of nickel sit harmlessly on the specimen surface.

Experimental spectra were recorded for the eight thermally oxidized specimens. Peak positions and energy differences will be discussed first from the standpoint of determining the possibility of structural ordering in the oxides. XPS thickness measurements will then be compared to the ellipsometrically determined values for these specimens. Lastly, peak intensity data will be presented, where variations of oxygen content in the films will be discussed.

Peak positions — The concept of structure-induced charge transfer (SICT) has been invoked (ref. 5 and references within) to relate small shifts in binding energy to variations in Si-O-Si bridging angles. Such variations can be described by the polymorphs (keatite, coesite, α-quartz, β-cristobalite) of SiO_2. We have evaluated the data from the thermally grown oxide films using the values of Grunthaner *et al.* (5).

Si2p peak differences (E_{Si+4} - E_{Si0}) for the thinnest sample, 3.7 nm, are 4.1 ± 0.1 eV over all five takeoff angles. The O1s signal is split into two peaks with energy differences of 430.5 ± 0.1 eV and 429.7 ± 0.2 eV relative to the Si^{+4} value. These values, especially those of silicon, are indicative of coesite and amorphous SiO_2 or α-quartz structures.

The O1s spectra for all eight experiments reveal a split oxygen peak with energy differences (E_B(O1s) - $E_B(Si^{+4})$) of 430.5 ± 0.1 eV and 429.6 ± 0.2 eV. The first peak matches the coesite energy difference of Grunthaner *et al.* while the latter defines either an amorphous SiO_2 or α-quartz structure. The possibility of the presence of the coesite structure has been

discussed (5, 7) as a short-range ordered component of the oxide near the silicon interface. Our evidence lends credence to this idea. During the very earliest stages of silicon oxidation, the growing oxide may be constrained by the orientation of atoms on the silicon surface. Such oxidation would necessarily occur epitaxially or quasi-epitaxially, until the oxide was sufficiently thick and/or continuous for oxide volume energy to overtake the surface energy dictated structure. Relaxation to a more thermodynamically stable structure could then proceed, with crystal volume changes accommodated by expansion of the film in the z-direction, normal to and away from the interface.

The relaxation of the oxide has been treated as a viscoelastic stress relief process by several authors (2, 8, 9). The *Anneal* protocol described in the previous section was designed to test the viscous flow model proposed by Irene *et al.* (2). The three annealed specimens presented silicon energy differences of 4.7 ± 0.1 eV and O1s differences of 430.5 ± 0.1 eV and 429.6 ± 0.2 eV, irrespective of final oxide thickness or XPS takeoff angle. These results are indistinguishable from those of the other samples, suggesting that viscous flow does not occur or occurs on a time scale much longer than previously reported.

The thinnest sample, a 3.7 nm film, exhibits a shift in the Si^{+4} peak as a function of takeoff angle as shown in figure 4. At grazing incidence, the Si2p energy difference ($E_B(Si^{+4})$ - $E_B(Si^0)$) is 4.6 eV, characteristic of α-quartz or keatite. The Si^{+4} peak shifts to 4.1 ± 0.1 eV for the higher angles, while complementary behavior is observed in the O1s double peak, suggesting a layered structure consisting of an outer layer of amorphous SiO_2 and an inner layer resembling coesite. The density of coesite is $\rho = 2.9$ g cm^{-3} (versus 2.20 g cm^{-3} for uncompacted vitreous silica). The existence of coesite at the interface may be explained by a quasi-epitaxial ordering of the SiO_2 film on the silicon substrate. This would lessen the interfacial stresses caused by lattice mismatch between the two. At greater

distances from the interface a pseudopolymorphic transformation to amorphous SiO_2 or α-quartz may occur. The presence of a high density interfacial layer would also increase the refractive index at the interface, consistent with the refractive index gradient reported in Part I.

Thickness calculations — Ellipsometric measurements were performed throughout the various stages of surface modification. High-speed polarization modulation ellipsometric data were collected as part of the kinetic study reported in Part III of this series. Two-zone null ellipsometry was used periodically to obtain data of high accuracy; only the latter data will be discussed here.

The null data obtained after oxide growth (and post treatment, where applicable) were input to McCrackin's ellipsometry program (10) to solve for both film thickness and index of refraction. Thicknesses were also calculated from Si^0 XPS signal attenuation measurements, where the Si^0 data from the *in situ* flash anneal provided a clean I_0 signal. The data show very good agreement (Table I) for six of the eight experiments; discrepancies arise when the film thickness falls below 5 nm.

Table I also contains the calculated film refractive index, n_f, which accompanies the calculated ellipsometric film thickness data. The trend of the indexes reflects that reported in Part I of this series, as well as the findings of Kalnitsky *et al.* (11). Meaningful solutions for n_f were found for the specimens of the *Anneal* protocol after film formation but prior to the anneal; meaningful solutions were not found after the anneal. We attribute this to interface roughening, possibly due to chemical mixing (the SiO_2/SiO phase boundary lies at 1150° C, see ref. 4). This should be a topic for further study.

Peak intensities — Calculations based on integrated peak intensities can provide elemental ratios within the sampled volume. While XPS is considered a surface sensitive technique, the surface sensitivity is not absolute (12). Both surface and near-surface regions can be analyzed by understanding the electron sampling depth and its dependence on analysis geometry. The sampling depth is conveniently defined by λ_m, the inelastic mean free path (IMFP) in monolayers.

λ_m represents a statistical signal attenuation length of the form 1/e. The escape depth can be defined by the relation (12) $d_{max} = 3\lambda_m$, where the proportionality constant of 3 assures that 95% of the total recorded electron fluence emanates from atoms of depth no greater than d_{max}. The depth of analysis is controlled by tilting the sample with respect to the axis of the photoelectron energy analyzer, resulting in the simple relation $d = 3\ \lambda_n \sin\theta$ where $\lambda_n = a\ \lambda_m$ is the IMFP in nanometers and θ is the angle between the sample surface and the axis of the detector optics. Table II shows our calculated depth values for the five takeoff angles used in this study.

Stoichiometric results for the eight thermal oxides are presented in Figure 5. A persistent and marked elevation in oxygen content is apparent in all samples. This observation could be an experimental artifact due to either instrument malfunction/miscalibration or sample contamination. To investigate these possibilities, a high-precision α-quartz resonator was cleaned in reagent grade acetone and subjected to a 30 second dip in 10% HF, rinsed in deionized water, blown dry and inserted into the XPS. Analysis was preformed under conditions similar to those of the thermal analyses ($\theta = 75°$ only). An O/Si^{+4} ratio of 2.08 was obtained. We attribute the slight oxygen enrichment to adsorbed species, most notably H_2O. The double oxygen peak is no longer present, and peak energy differences are consistent with those of the a-quartz structure.

The thermal oxide stoichiometry data are replotted in Figure 6 to show the variation in oxygen content with distance from the substrate/oxide interface. The negative values indicate complete penetration of the oxide with an increasing Si^0 signal component from the substrate. The casual observer may be tempted to divide the data into two natural groups: "thin" films that can be analyzed throughout their depth and "thick" films in which the electron escape depths barely equal the film thickness. This would lead to the erroneous conclusion that the thin films contain higher quantities of oxygen, when in fact adsorbed species make up a major portion of the O1s signal. The contribution of the surface layers may be removed by the approximation

$$x_{subsurface} = [(x_{film} \cdot d_{film}) - (x_{surface} \cdot d_{surface})]/d_{subsurface} \quad [4]$$

This calculation results in a mean O/Si^{+4} ratio of 2.67 ± 0.16 for all samples.

A dissolved oxygen concentration of ~5×10^{21} cm^{-3} would be required to produce this signal, a value five orders of magnitude greater than the equilibrium concentration for dissolved oxygen in SiO_2 of 5×10^{16} cm-3. Han and Helms (14) reported similar values in an ^{18}O tracer study of oxygen transport through the oxide. They speculated that the high concentration (which was found on the outer surface of the oxide) could not be an adsorbed species, but followed a concentration gradient in the outer 10 nm of oxide. From their data, they calculated a Debye length at 1000° C of $L_D = 2.7$ nm. This value corresponds roughly to the width of the refractive index gradient reported in Part I of this paper. The presence of a coesite-like structure at the interface may provide a transition zone for accommodation of the lattice mismatch between silicon and the growing SiO_2, while a large concentration of defects about the interface (e. g. dislocations, mixed 4- and 6-membered

ring structures) may provide easy pathways for highly enhanced oxygen solubility in the oxide.

Conclusions

1. XPS data from ultrathin SiO_2 films thermally grown on silicon (111) substrates indicate a mixture of the SiO_2 polymorphs coesite (a 4-membered ring structure) and amorphous SiO_2 or α-quartz, with a preponderance of coesite near the film-substrate interface.

2. The existence of coesite (a high density polymorph) at the interface would account for the refractive index gradient reported in Part I of this series and for the high refractive indexes reported for the films used in this study.

3. A high temperature anneal was performed to test the viscoelastic stress relief hypothesis. No chemical differences were found between annealed and unannealed samples. A loss of optical properties did occur, which we attribute to interfacial roughening due to chemical mixing.

Part III of this series will examine the kinetics of the initial silicon oxidation reaction.

Acknowledgments

This work was supported by the Office of Naval Research under grant no. N00014-89-J-1265.

References

1. G. A. Danko, J. Kruger, and A. G. Revesz, "Rapid Film Formation on Clean Silicon Surfaces I: Optical Properties", *to be submitted.*

2. E. A. Irene, E. Tierney, and J. Angilello, *J. Electrochem. Soc.*, **129**, 2594 (1982).

3. E. A. Gulbransen and S. A. Jansson, *Oxid. Met.*, **4**, 181 (1972).

4. G. A. Danko, "A System for the Study of the Growth of Silicon Oxide Films with Real-Time Process Monitoring Capability", *to be submitted.*

5. F. J. Grunthaner, P. J. Grunthaner, R. P. Vasquez, B. F. Lewis, J. Maserjian, and A. Madhukar, *J. Vac. Sci. Technol.*, **16**, 1443 (1979).

6. V. D. Borman, E. P. Gusev, Yu. Yu. Lebedinskii, and V. I. Troyan, *Phys. Rev. Lett.*, **67**, 2387 (1991).

7. B. J. Mrstik, A. G. Revesz, M. Ancona, and H. L. Hughes, *J. Electrochem. Soc.*, **134**, 2020 (1987).

8. E. P. EerNisse, *Appl. Phys. Lett.*, **35**, 8 (1979).

9. J. R. Patel and N. Kato, *J. Appl. Phys.*, **44**, 971 (1973).

10. F. L. McCrackin, NBS Technical Note 479, U. S. Government Printing Office, Washington, DC (1969).

11. A. Kalnitsky, S. P. Tay, J. P. Ellul, S. Chongsawangvirod, J. W. Andrews, and E. A. Irene, *J. Electrochem. Soc.*, **137**, 234 (1990).

12. D. Briggs and M. P. Seah, eds. Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy. John Wiley & Sons, New York (1983).

13. M. P. Seah and W. A. Dench, *Surf. and Interf. Analys.*, **1**, 2 (1979).

14. C.-J. Han and C. R. Helms, *J. Electrochem. Soc.*, **135**, 1824 (1988).

Sample	d_{XPS}	d_{ellips}	n_f
Flox.4	1.28	5.26	3.190
Anneal.6		4.24[a]	2.325
	1.30	4.72[b]	
Anneal.2		4.29[a]	2.063
	4.43	5.49[b]	
Flox.6	5.85	6.13	2.009
Anneal.4		6.42[a]	1.845
	8.17	8.09[b]	
Noflox.2	8.07	8.00	1.409
Flox.2	9.72	9.12	1.564
Noflox.4	9.22	9.26	1.420

[a]Before anneal. [b]After anneal, with n_f forced to 1.461.

Table I. Film thicknesses as measured by ellipsometry and XPS. All ellipsometric measurements were made at a wavelength of 632.8 nm. Thickness values are given in nanometers.

θ	$\lambda_n(Si0)$	$\lambda_n(Si^{+4})$	$\lambda_n(O)$
15	2.82	2.82	2.24
30	5.45	5.44	4.32
45	7.71	7.70	6.11
60	9.44	9.43	7.48
75	10.5	10.5	8.34

Table II. Spectrometer takeoff angles and depth calculations used in this work. Inelastic mean free paths are given in nanometers. See text for details.

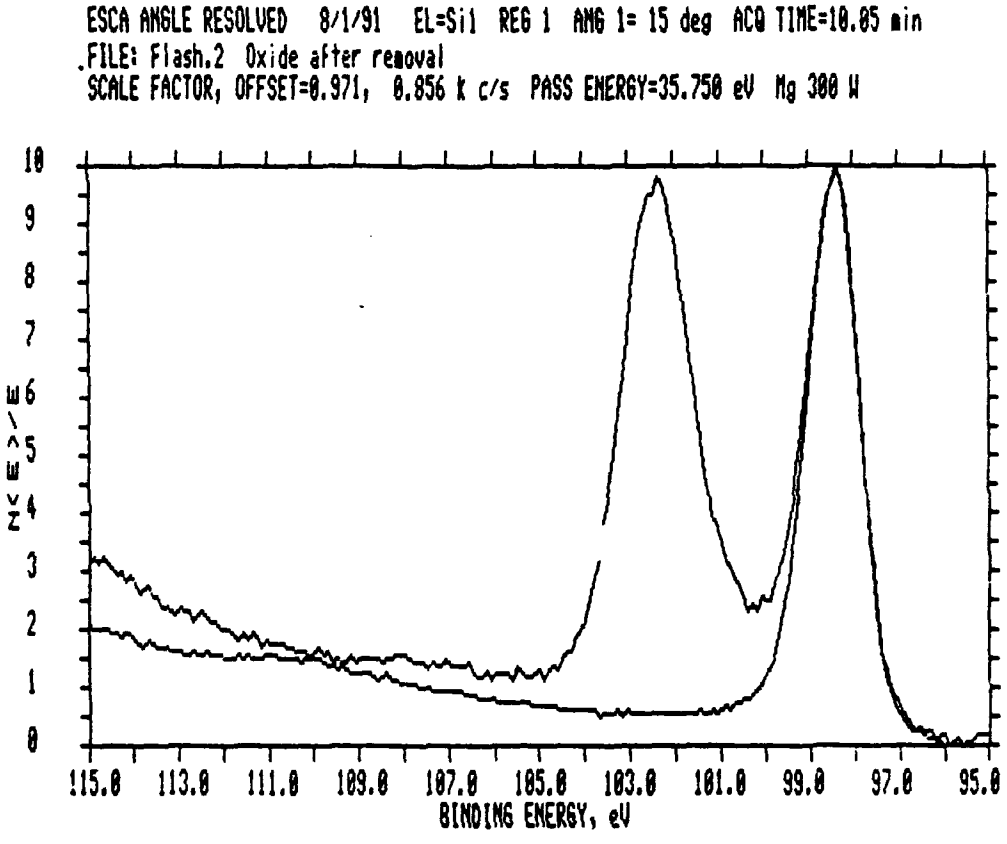


Figure 1. (a) $Si2p_{1/2}$ x-ray photoelectron spectrum. The absence of the peak at 103 eV is clear evidence of oxide removal.

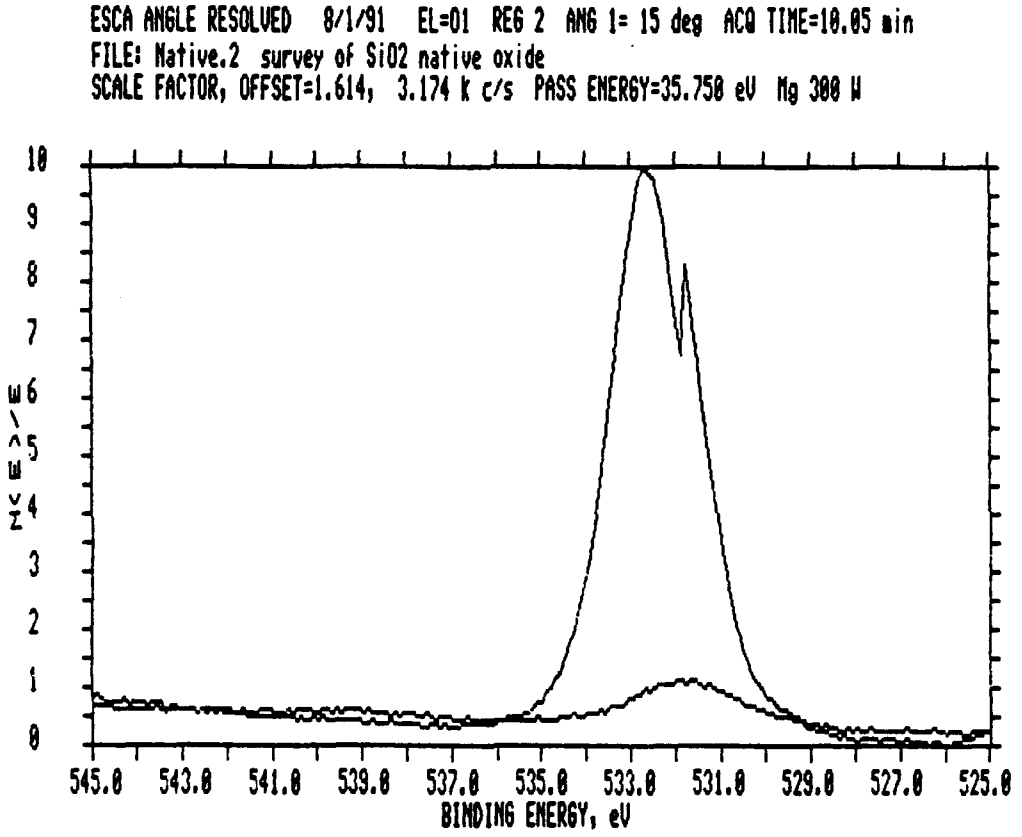


Figure 1. (b) O1s spectrum. Oxygen removal is not complete, but the remaining signal accounts for only about 8 atomic percent of the material in the top 2.2 nm of the sample.

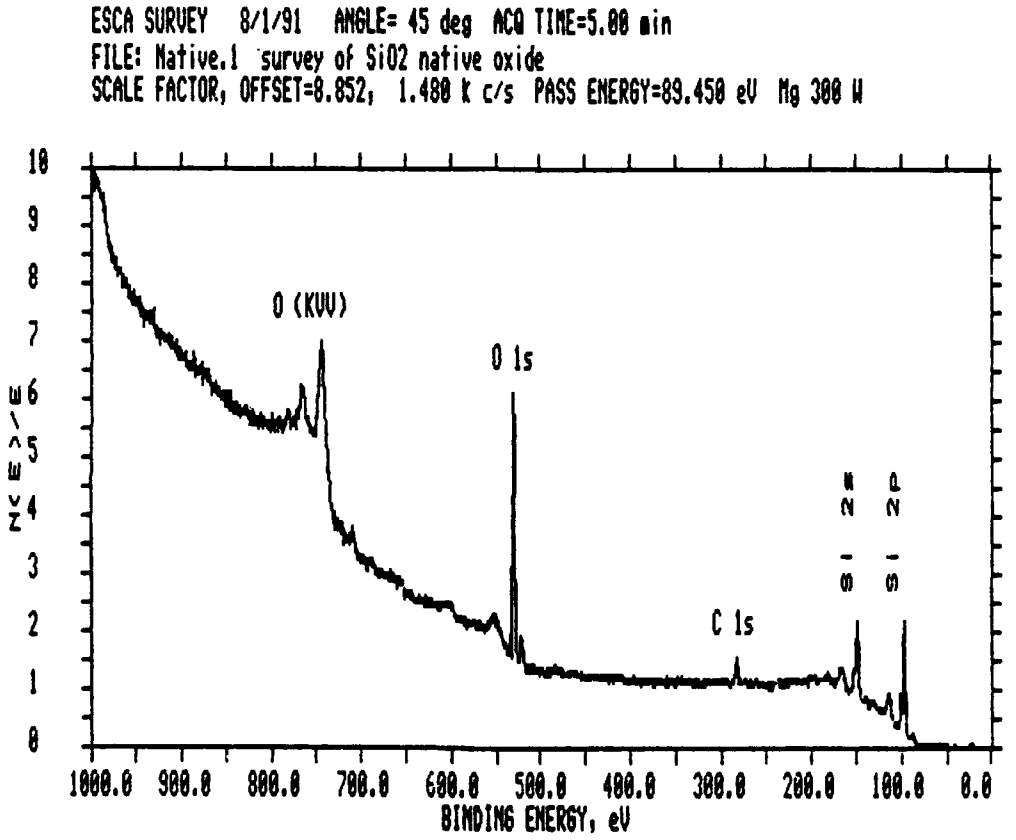


Figure 2. (a) Survey spectrum of a native oxide on silicon.

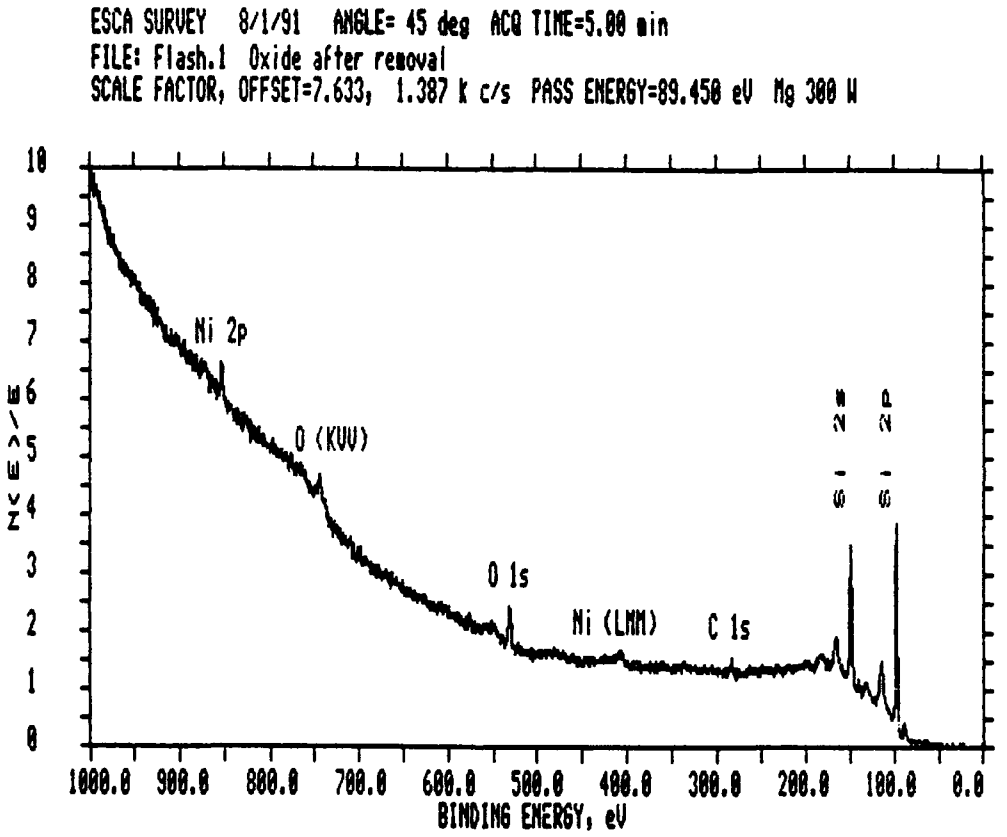


Figure 2. (b) After the flash anneal, the oxygen signal is suppressed, while a small quantity of nickel has appeared. This was found to be due to ablation of the Inconel grips.

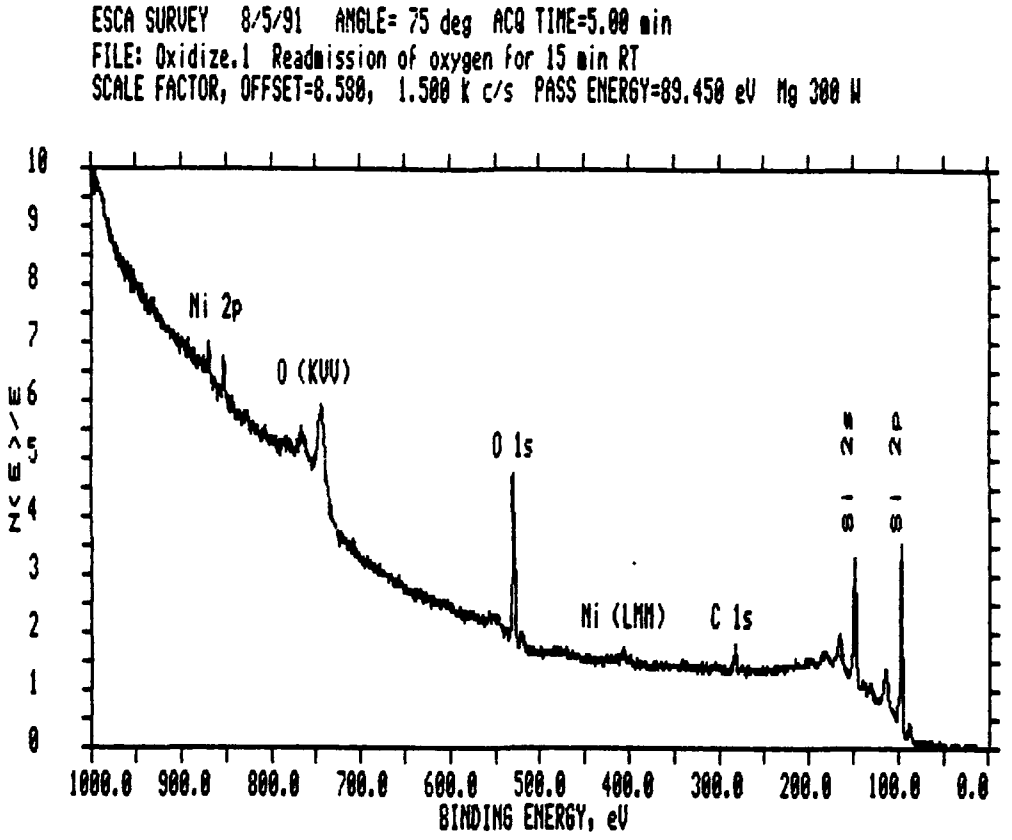


Figure 2. (c) Readmission of oxygen at room temperature results in reformation of the native oxide.

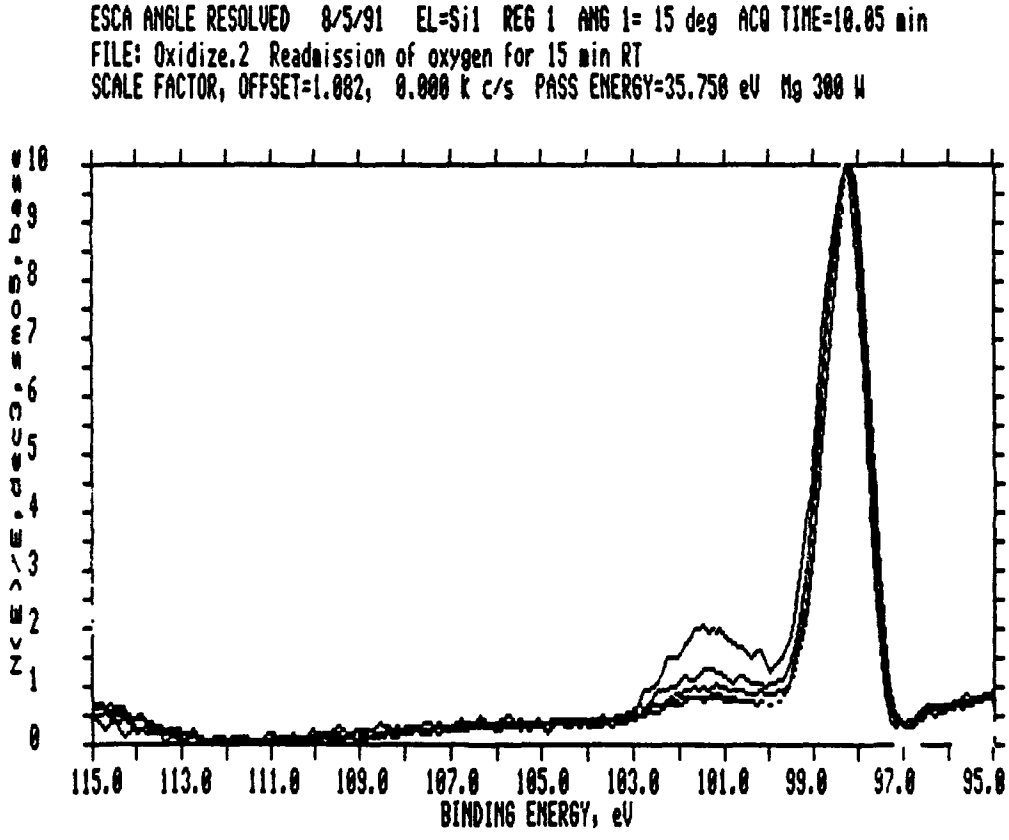


Figure 3. (a) Si2p spectrum of the reformed native oxide, revealing valence states less than +4.

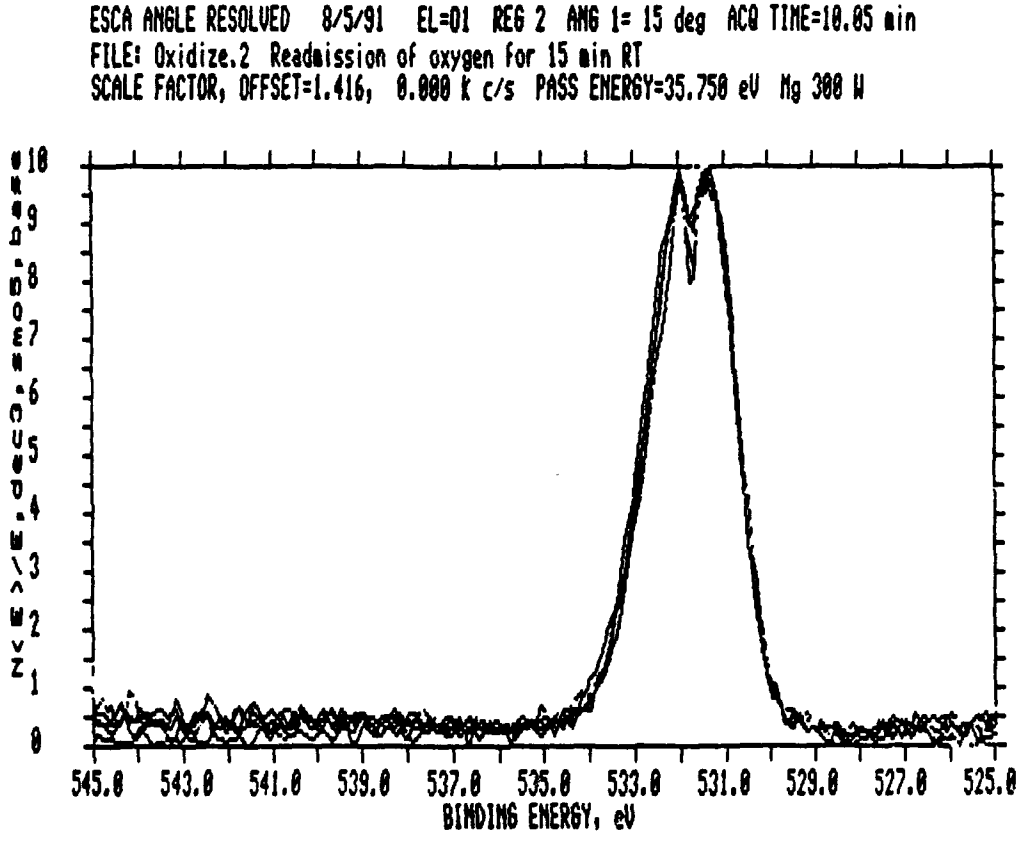


Figure 3. (b) O1s spectrum. A peak split is evident, resulting from structure-induced charge transfer. See text for details.

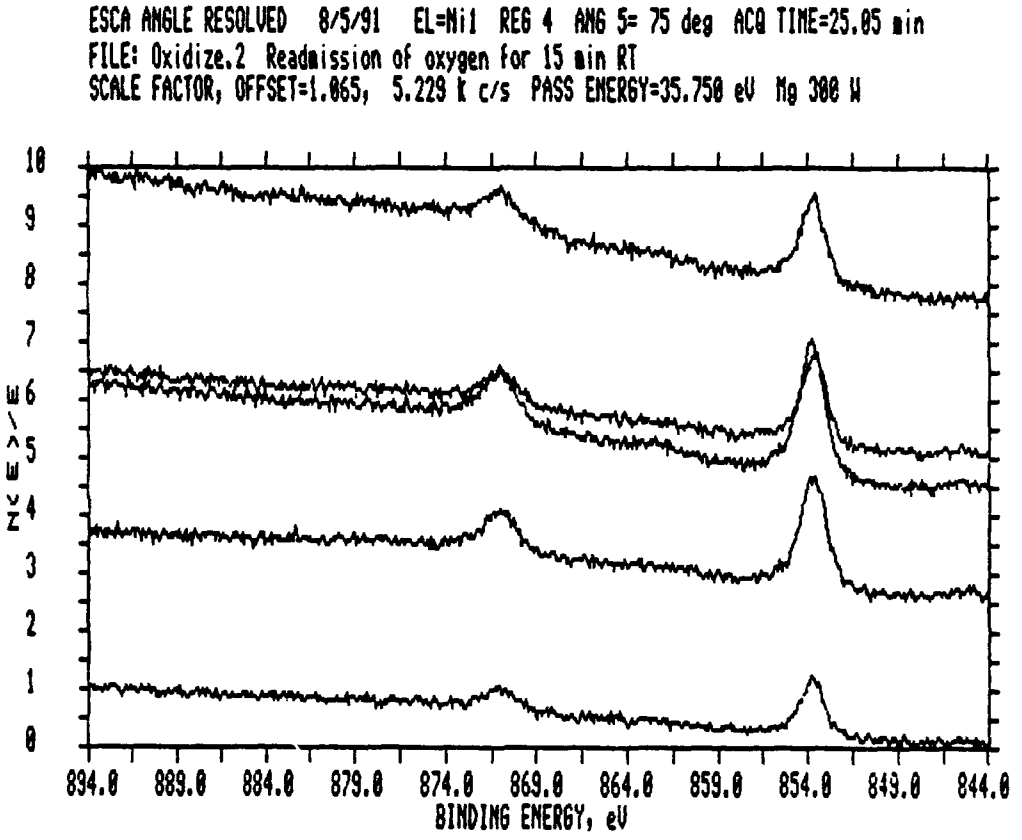


Figure 3. (c) Ni2p doublet after admission of oxygen. The invariance of the peaks suggests the presence of a few large particles of metallic nickel on the specimen surface.

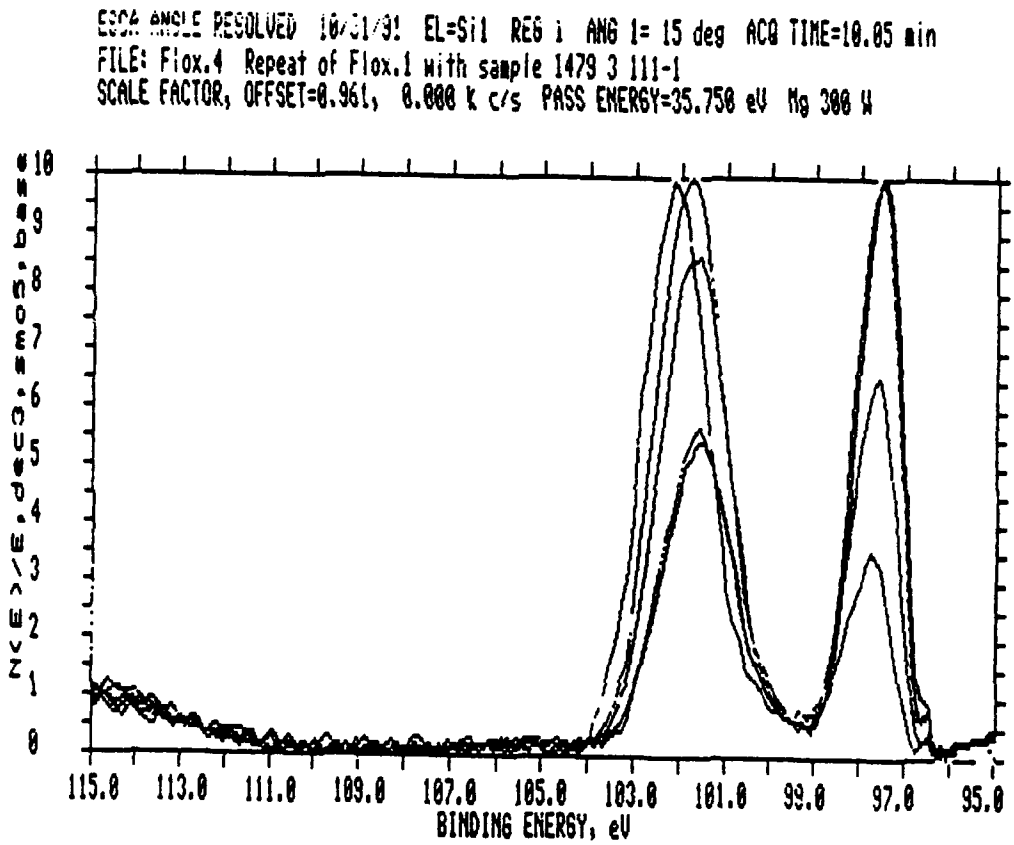


Figure 4. Si2p spectrum of the thinnest (3 nm) SiO_2 film. The Si^{+4} peaks shift from right to left with increasing takeoff angle, suggesting that coesite-like 4-membered ring structures exist preferentially at the bottom of the oxide, i. e. near the Si/SiO_2 interface.

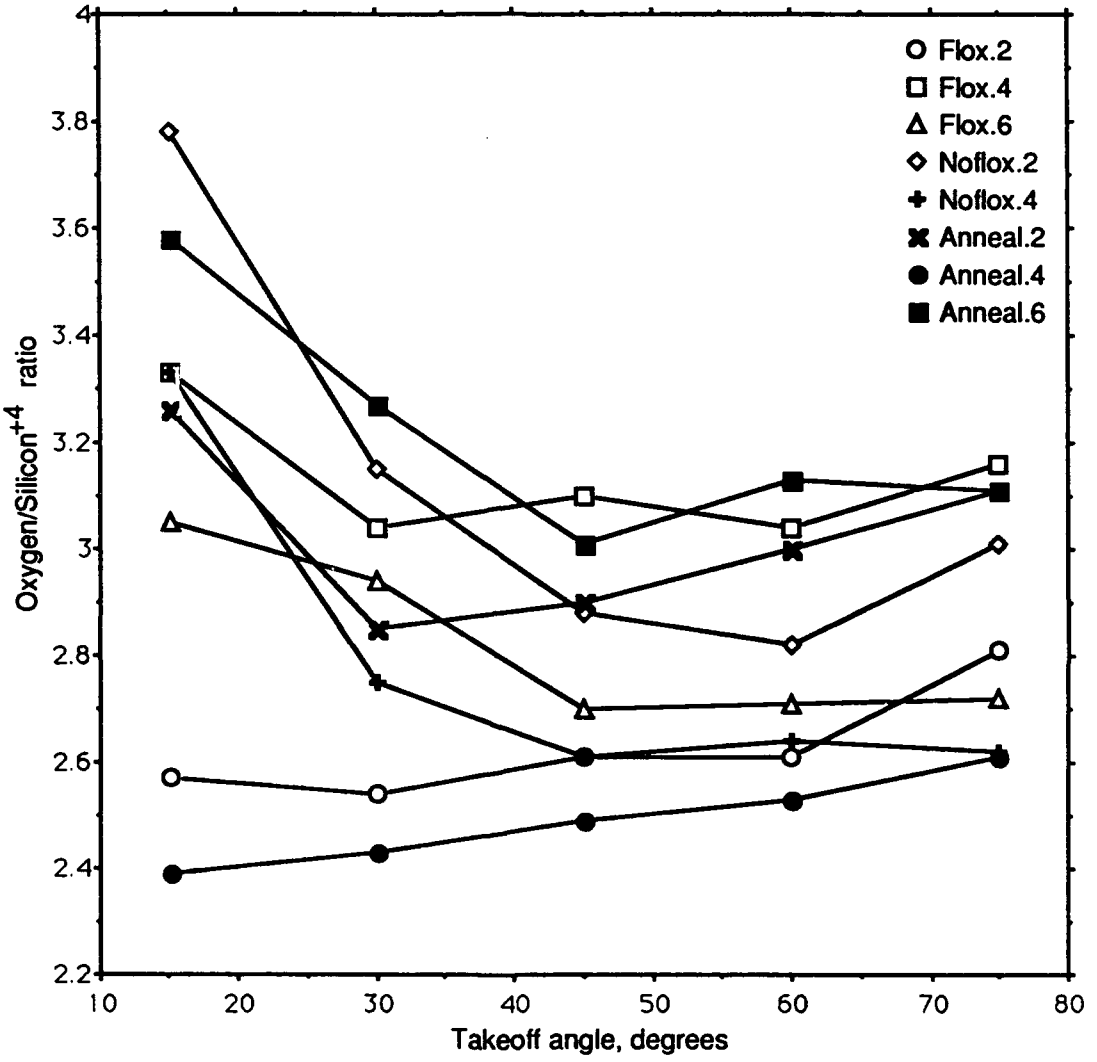


Figure 5. Stoichiometry of thermal oxides vs. takeoff angle. Note the persistent supersaturation of oxygen in all samples. The high O/Si^{+4} ratios at grazing angles are due to surface contamination, but the magnitude of the signal is insufficient to account for the elevation at the higher takeoff angles.

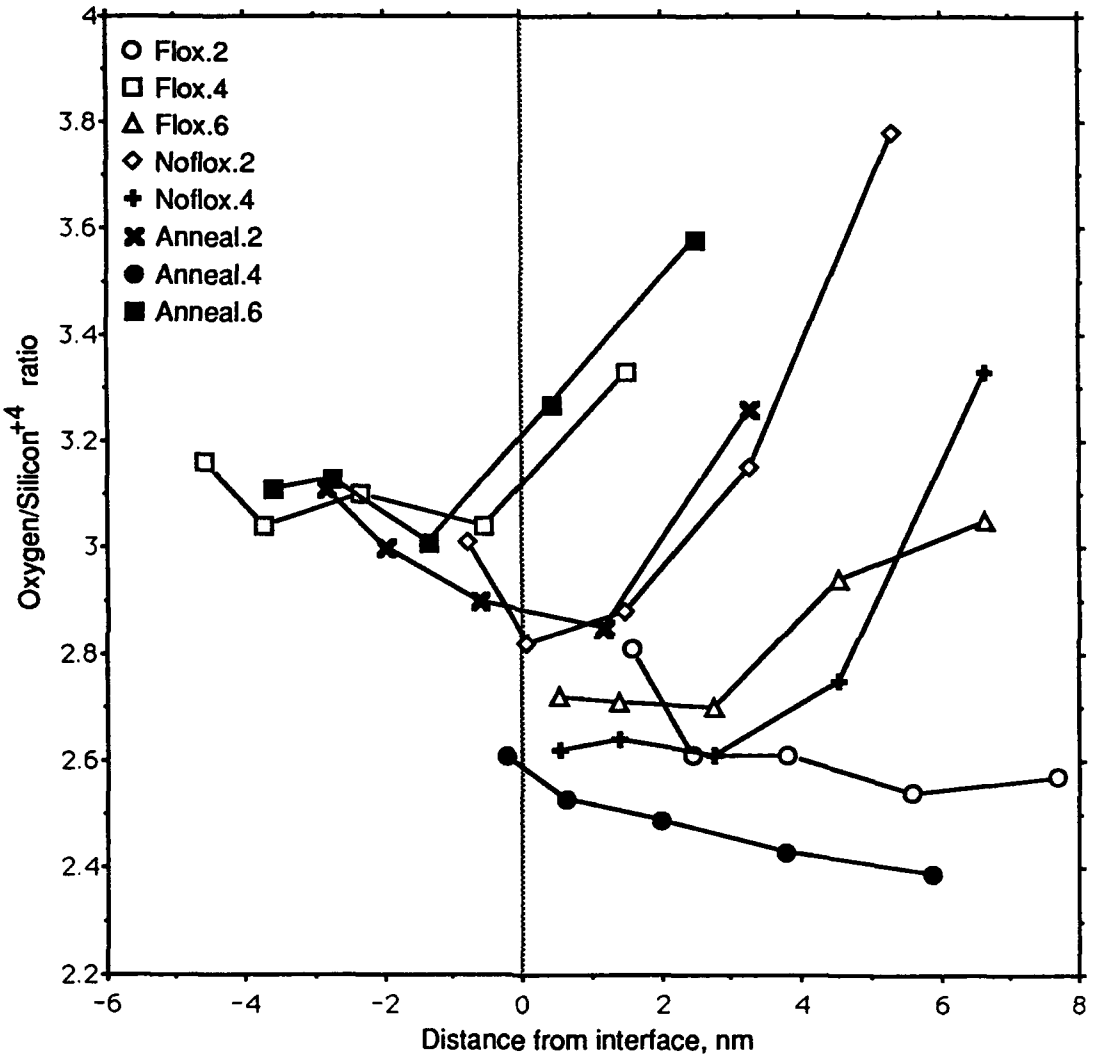


Figure 6. Data from figure 5, replotted to account for sampling depth. Each point on the graph represents the maximum sampling depth $d = 3 \lambda_n \sin \theta$. Negative values indicate that the entire oxide is analyzed, as well as some of the underlying substrate.

Section V

Kinetics of SiO_2 Film Growth

A paper to be submitted to the *Journal of the Electrochemical Society*

Rapid Silicon Dioxide Film Formation on Clean Silicon Surfaces III: Kinetics

G. A. Danko, J. Kruger

Department of Materials Science and Engineering

The Johns Hopkins University, Baltimore, Maryland 21218

A. G. Revesz

Revesz Associates, Bethesda, Maryland 20817

ABSTRACT

Real-time *in situ* ellipsometric measurements of silicon oxidation were carried out. Flash annealed samples for three orientations were exposed to dry oxygen at temperatures from 800° C to 1000° C and the reaction followed by polarization modulation ellipsometry. Two new linear oxidation regimes are observed for the initial stages of film formation. We report activation energies for these reactions, discuss the limiting film thicknesses between the regimes, and discuss the impact of these new regimes on the refractive index of thin SiO_2 films.

Introduction

Silicon oxidation time-thickness relationships have been the subject of a large research effort since before the general formulation of Deal and Grove (1). Much of the information

published deals with possible mechanisms of oxide growth (2-8) and theoretical formulations (8-11). These efforts have dealt with behavior in both the linear and parabolic regimes of Deal and Grove, as have many experimental studies (8, 12-14). Without exception, these investigators are unable to extrapolate their data or their models back to thickness d = 0 at time t = 0; a limiting thickness of approximately 3 nm has prevented study of thinner oxides. Tiller (15) has suggested a blocking layer based on interfacial stresses which inhibits further progression of an initial rapid reaction, while Paz de Araujo *et al.* (16) (after Murali and Murarka (17)) has discussed the possibility of an oxygen-rich subsurface layer within the silicon substrate which provides for an enhancement of an initial interfacial reaction.

We report the results of experimental work in which the time-thickness profile of silicon oxidation was recorded by ellipsometric observation. Flash annealed silicon was brought to process temperature *in vacuo* and oxygen was introduced, permitting growth measurements from zero film thickness to any desired endpoint. Two new linear oxidation regions have been observed and their activation energies have been calculated. We refer to these regions as Region I and Region II, thus the established Deal and Grove linear region could be termed Region III, with Region IV representing parabolic growth.

The transition between Regions I and II occurs as a function of reaction temperature and may explain the blocking layer phenomenon discussed above. We refer to this layer as a limiting film thickness. It will be related to a graded refractive index in the oxide reported in Part I of this paper and to oxygen supersaturation and structure-induced charge transfer effects reported in Part II.

Experimental Procedures

2-inch diameter silicon wafers of orientations (100), (111), and (110) were procured from Virginia Semiconductor, Inc. As received, wafers were B-doped (p-type) 0.1 ohm-cm and polished on one side only. 5 x 20 mm samples were cleaved, edge polished, and RCA cleaned as described previously (18). The samples were cleaned in small batches and stored under clean room conditions prior to use; in no case did the storage period exceed ten days.

The growth chamber and polarization modulation ellipsometer (PME) are described in detail elsewhere (19). The growth chamber is capable of pressures of 10^{-8} torr. Such low pressures are required for the final specimen cleaning procedure, a "flash anneal" which sublimes the native oxide from the substrate, leaving an essentially oxide-free silicon surface. Sample heating was effected ohmically by passing current-limited DC through the substrates supported in Pt-coated Inconel grips. Temperatures up to the melting point of silicon were readily achieved.

Sample temperatures were measured by optical pyrometry (λ = 650 nm). Suitable corrections were made for sample emissivity and loss of light intensity due to the needed mirrors and window in the imaging path. Accurate temperature determination is crucial both to the proper selection of substrate optical constants and as an independent variable for the kinetic study that is the topic of this paper.

Ellipsometric observation was chosen to follow film growth because it is both non-invasive and nondestructive. For this project, we built a polarization modulation ellipsometer capable of data rates of 1000 sec^{-1}. The drawbacks of the PME are its limited accuracy (drift) and poor signal to noise ratio. The instrument is thus unsuitable for absolute

measurements, but ideal for relative measurements at high speed. Occasional two-zone null ellipsometry measurements provided the high accuracy data to which the PME data could be corrected. This was especially important when assessing the optical properties of the hot silicon surface immediately prior to oxygen introduction.

Our experimental approach was patterned after that of Hopper, Clarke, and Young (20). RCA cleaned samples were individually loaded into the growth chamber, evacuated to 10^{-8} torr, and subjected to a 1200° C flash anneal for 15 seconds. At 10^{-8} torr, monolayer-scale contamination occurs in approximately 100 seconds; this is a sufficient time interval to obtain a manual two-zone ellipsometric measurement, establish a process temperature, begin PME data acquisition at 10 points sec^{-1}, and introduce ultrahigh purity Matheson oxygen. Oxidation of three silicon surface orientations was investigated primarily at five temperatures from 800° C to 1000° C. Data from one experiment at 913° C (due to an improper pyrometer setting) is also included. All experiments were performed at 1 atm pressure in a dry oxygen ambient. Oxidation times varied from 20 minutes to over three days (one ellipsometric period). A total of 27 experiments were performed.

Results and Discussion

Data from a typical experiment are presented in Figure 1. Three distinct regions are observed. Region 0 represents the condition of the sample surface before oxygen is admitted to the chamber. This region provides baseline data Δ, Ψ; the reader is directed to Part I of this paper for further details. Region I, when dry O_2 is admitted to the chamber, reveals rapid film growth by the distinct shift in Δ. Region II is the start of the second new linear growth region.

The initial stages of thermal oxidation are of interest as the size of microelectronics structures continues to shrink. We have investigated this behavior, to which we refer as Region I, and will discuss the activation energy of this process and offer predictions of the minimum achievable oxide film thickness as a function of reaction temperature. Region II has been measured and an activation energy determined, as we report below.

Region I — The initial oxidation of silicon is achieved in our experiments by flooding the growth chamber with dry ultrahigh purity Matheson oxygen. One atmosphere of pressure is reached in approximately 5 seconds, hence the active-to-passive oxidation threshold (21, 22) is crossed within a few milliseconds of the opening of the gas valve. Film formation proceeds rapidly until a limiting thickness is reached, at which time the rapid growth abruptly slows to the linear Region II behavior. Two possible oxidation mechanisms may be operative: 1) the formation of islands and preferential oxidation along island edges, resulting in lateral two-dimensional island growth, or 2) random sticking of oxygen molecules and highly localized, uncoordinated oxidation which results in uniform film thickening. The first case can involve competitive active and passive oxidation as observed by Smith and Ghidini (23); this mechanism would result in a rough (~2 nm) interface between substrate and oxide as seen by Helms *et al.* (24). Such behavior will severely restrict the validity of ellipsometric observations prior to island coalescence. Even effective medium modelling techniques (25, 26) are of limited value without first-hand knowledge of island thickness, refractive index, and two-dimensional spatial distribution. The second case, typified by recent scanning tunneling microscopy studies (27), offers the possibility of valid ellipsometric observation. Whatever the mechanism, this reaction is driven until a limiting film thickness is attained that terminates rapid film growth. Let us consider reaction rate and termination in turn.

In Figure 2 we present the experimental data as $\ln \frac{dL}{dt}$ vs. $\frac{1}{T_{abs}}$. The scatter in the data reflects the inherent difficulties in an experiment of this type. We have fit the data by regression analysis (n = 27) to obtain values for an Arrhenius-type expression. The rate constant for the oxide in Region I is

$$\frac{dL}{dt} = 5.8 \cdot \exp\left[- \frac{0.603 \text{ eV}}{kT} \right] \text{nm sec}^{-1} \qquad [1$$

The activation energy, 0.603 eV, when restated in calories per mole, is E_I = 13.9 kcal mol^{-1}. This value compares very favorably with that obtained by Gelain *et al.* (22) (E_A = 13 ± 1 kcal mol^{-1}) for their active-passive oxidation boundary studies (we note that Gelain *et al.* conducted a thermodynamic measurement, and no rate studies were carried out). The value they quote is ascribed to the active oxidation or "combustion" reaction of silicon; it is a reaction with kinetics of the first order. The high oxygen pressures in our growth chamber (several torr, on the time scale of this behavior) retard SiO movement away from the surface; if a SiO fog can form it immediately reprecipitates on the surface, leading to the formation of a continuous and protective SiO_2 film.

Limiting film thicknesses are shown in Figure 3. The individual data points were obtained from straight-line extrapolation of the Region II Δ, Ψ values as typified by Figure 1. The loci of intersection of the Δ, Ψ lines with the time value of the upper inflection point of the Δ curve (the start of Region I) is used to specify the starting values for Δ and Ψ. These values are corrected to the more accurate manually-acquired null ellipsometric values, defined as Δ and Ψ. The data obtained similarly from the lower inflection point of the Δ curve (the time value where Region I and Region II meet) provides Δ_f, Ψ_f which are corrected by a like amount. Thickness and refractive index values could then be solved by

computer modelling using McCrackin's ellipsometry program (28). Substrate optical constants were chosen based on the data of van der Meulen and Hien (29).

This method produces very coarse data, but in light of the very rapid and minute changes taking place on the substrate surface this procedure yields data of surprisingly good quality. Attempts to assign a functional relationship to these data have not been successful; linear and polynomial regressions result in $r^2 = 0.427$ and $r^2 = 0.46$, respectively. Smith and Ghidini (23) report a limiting thickness of s = 7.2 nm based on their competitive kinetics model. A linear extrapolation of our data to 1100° C results in s = 3.2 nm, while a second order polynomial results in s = 12.2 nm. These values are within a factor of two of that of Smith and Ghidini, suggesting consistency with their model.

In Part II of this paper, we reported a large oxygen supersaturation in our films. Han and Helms (30) report a supersaturation gradient at the outer oxide surface in their ^{18}O tracer studies; they calculated a Debye length of $L_D = 2.7$ nm to support the large solubility value observed ($n = 5 \times 10^{21}$ cm^{-3}). We have seen similar oxygen levels in films less than 10 nm in thickness, and speculate that Region I growth may be a field-assisted phenomenon, as originally proposed by Deal and Grove. The refractive index gradient reported in Part I may also be explained by this region, as Paz de Araujo *et al.* allude to quantities of unreacted subsurface oxygen and a partially reacted silicon overlayer.

Region II — The initial burst of oxidation is overtaken by a slower linear mechanism. We have computed a rate constant for this reaction,

$$\frac{dL}{dt} = 77 \cdot \exp\left[- \frac{0.794 \text{ eV}}{kT} \right] \text{nm sec}^{-1} \qquad [2]$$

This linear region seems to hold for oxide thicknesses between 1 nm and 3 nm and is operative for about two minutes at 800° C. As reaction temperature increases, this reaction is operative for shorter times, again apparently subject to some limiting thickness criterion. The rapid thermal oxidation (RTO) model of Paz de Araujo *et al.* again may explain this behavior. Their RTO decay times (~100 sec) agree with the time scale of this phenomenon. This region does not have an easily identifiable end point; it melds into the conventional linear slope (Region III) at about 3-5 nm, where our data then agree closely with the surface reaction rate constants of Han and Helms (30). It is clear that this region is in a state of disorder; XPS data indicate that the short-range order that does develop could be a quantity of 4-membered rings similar to coesite, a high density phase of SiO_2 with $\rho = 2.9$ g cm^{-3}. The decay process may be an ordering of the interfacial region brought about by the formation of structure within the oxide. As the film starts to thicken (in the first fraction of a second), amorphous SiO_2 forms, with a coesite polymorph present at the interface to accommodate the lattice mismatch with the underlying silicon substrate (7). The decay time of Region II is the time required to establish steady-state interface-controlled SiO_2 film growth.

Once steady-state interface-controlled growth is established, the reaction is now fully in Region III, the accepted linear region of the linear-parabolic model. The data in the range of 3-10 nm agrees quite closely with the surface reaction rate constants of Han and Helms (30). Since this work was concerned with only the rapid growth of SiO_2, very few experiments were carried beyond 15 nm. The parabolic region was not studied extensively and data beyond 15 nm is too sparse to obtain the Region IV rate constant.

Summary and Conclusions

We have identified a two-step process for the initial formation of SiO_2 on silicon. This redefines the linear-parabolic growth model as a four-region process of initial film formation (Region I), two linear growth regions (II and III), and the parabolic film thickening region (IV).

1. Region I is a rapid combustion reaction. The reaction terminates abruptly at a limiting film thickness.

2. The limiting film thickness is temperature dependent,.

3. Region II growth exhibits linear behavior. It exists only for short times, supporting the rapid thermal oxidation (RTO) model of decay of the Region I reaction.

4. The growth behavior can be explained by optical constant and compositional and structural results which found a refractive index gradient and a coesite-like interfacial layer.

Acknowledgments

This work was supported by the Office of Naval Research under grant no. N00014-89-J-1265.

References

1. B. E. Deal and A. S. Grove, J. *Appl. Phys.*, **36**, 3770 (1965).

2. W. A. Tiller, *J. Electrochem. Soc.*, **128**, 689 (1981).

3. E. A. Irene, E. Tierney, and J. Angilello, *ibid.*, **129**, 2594 (1982).

4. E. A. Irene, *J. Appl. Phys.*, **54**, 5416 (1983).

5. C.-J. Han and C. R. Helms, *J. Electrochem. Soc.*, **134**, 1297 (1987).

6. R. H. Doremus, *ibid.*, **134**, 2001 (1987).

7. B. J. Mrstik, A. G. Revesz, M. Ancona, and H. L. Hughes, *ibid.*, **134**, 2020 (1987).

8. H. Z. Massoud, J. D. Plummer, and E. A. Irene, *ibid.*, **132**, 2685 (1985).

9. R. Ghez and Y. J. van der Meulen, *ibid.*. **119**, 1100 (1972).

10. W. A. Tiller, *ibid.*, **127**, 619 (1980).

11. W. A. Tiller, *ibid.*, **127**, 625 (1980).

12. Y. J. van der Meulen, *ibid.*, **119**, 530 (1972).

13. F. P. Fehlner, *ibid.*, **119**, 1723 (1972).

13. D. W. Hess and B. E. Deal, *ibid.*, **122**, 579 (1975).

14. W. A. Tiller, *ibid.*, **130**, 501 (1983).

15. C. W. Paz de Araujo, R. W. Gallegos, and Y. P. Huang, *ibid.*, **136**, 2673 (1989).

17. V. Murali and S. P. Murarka, *J. Appl. Phys.*, **60**, 2106 (1986).

18. G. A. Danko, J. Kruger, and A. G. Revesz, "Rapid Film Formation on Clean Silicon Surfaces I: Optical Properties", *to be submitted.*

19. G. A. Danko, "A System for the Study of the Growth of Silicon Oxide Films with Real-Time Process Monitoring Capability", *to be submitted.*

20. M. A. Hopper, R. A. Clarke, and L. Young, *J. Electrochem. Soc.*, **122**, 1216 (1975).

21. E. A. Gulbransen and S. A. Jansson, *Oxid. Met.*, **4**, 181 (1972).

22. C. Gelain, A. Cassuto, and P. Le Goff, *ibid.*, **3**, 139 (1971).

23. F. W. Smith and G. Ghidini, *J. Electrochem. Soc.*, **129**, 1300 (1982).

24. C. R. Helms, Y. E. Strausser, and W. E. Spicer, *Appl. Phys. Lett.*, **33**, 767 (1978).

25. P. J. McMarr and J. R. Blanco, *Appl. Opt.*, **27**, 4265 (1988).

26. D. E. Aspnes and J. B. Theeten, *Phys. Rev. B*, **20**, 3292 (1979).

27. Ph. Avouris, I.-W. Lyo, and F. Bozso, *J. Vac. Sci. Technol.*, **9**, 424 (1991).

28. F. L. McCrackin, NBS Technical Note 479, U. S. Government Printing Office, Washington, DC (1969).

29. Y. J. van der Meulen and N. C. Hien, *J. Opt. Soc. Am.*, **64**, 804 (1974).

30. C.-J. Tan and C. R. Helms, *J. Electrochem. Soc.*, **135**, 1824 (1988).

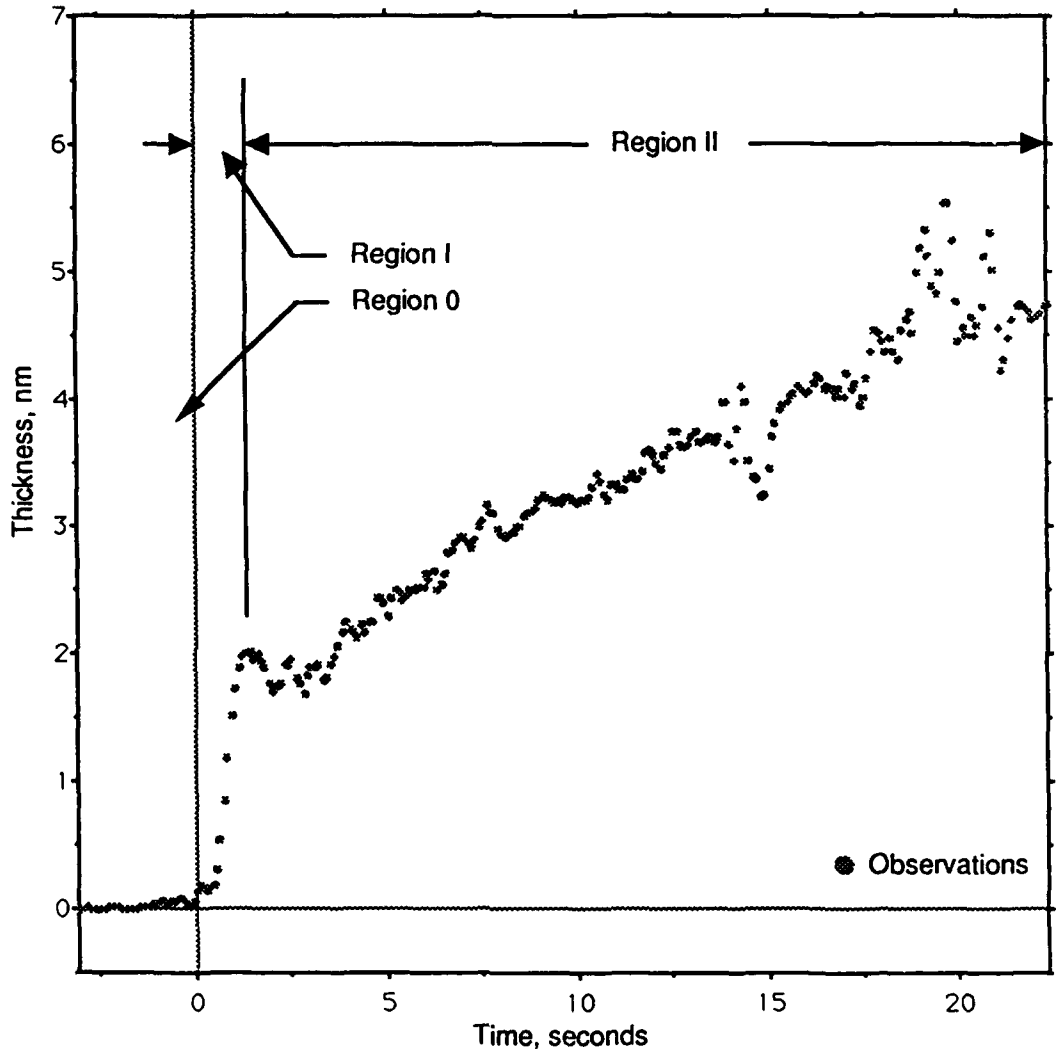


Figure 1. Time-thickness data from an experiment at 1000° C. Region 0 is the zero thickness baseline. In Region I, the rapid oxidation reaction is seen. Region II is a linear film growth region with a growth rate greater than the linear portion of the Deal and Grove model.

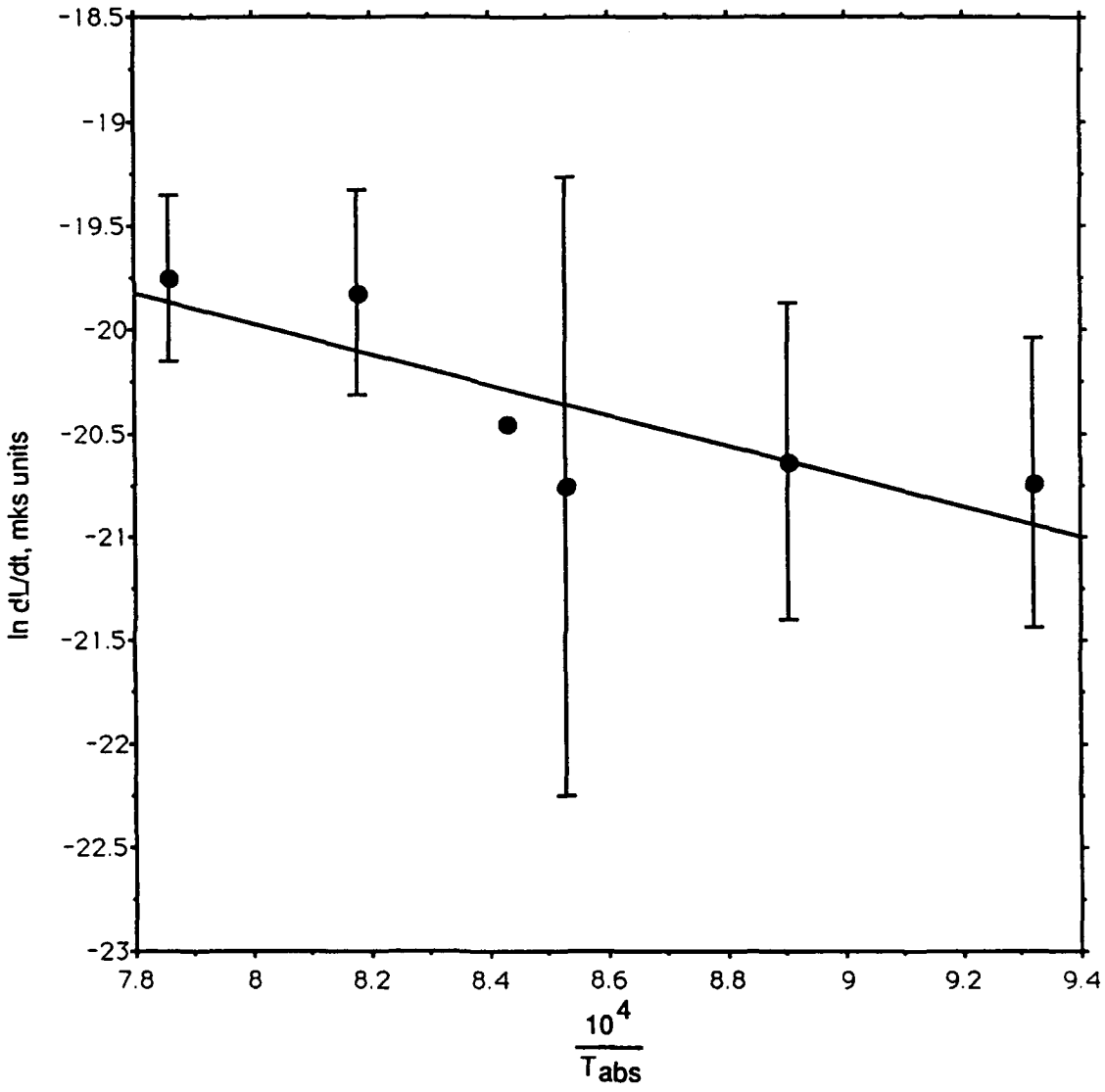


Figure 2. $\ln\left(\frac{dL}{dt}\right)$ vs. $1/T_{abs}$ for the Region I reaction. Data are presented in mks units.

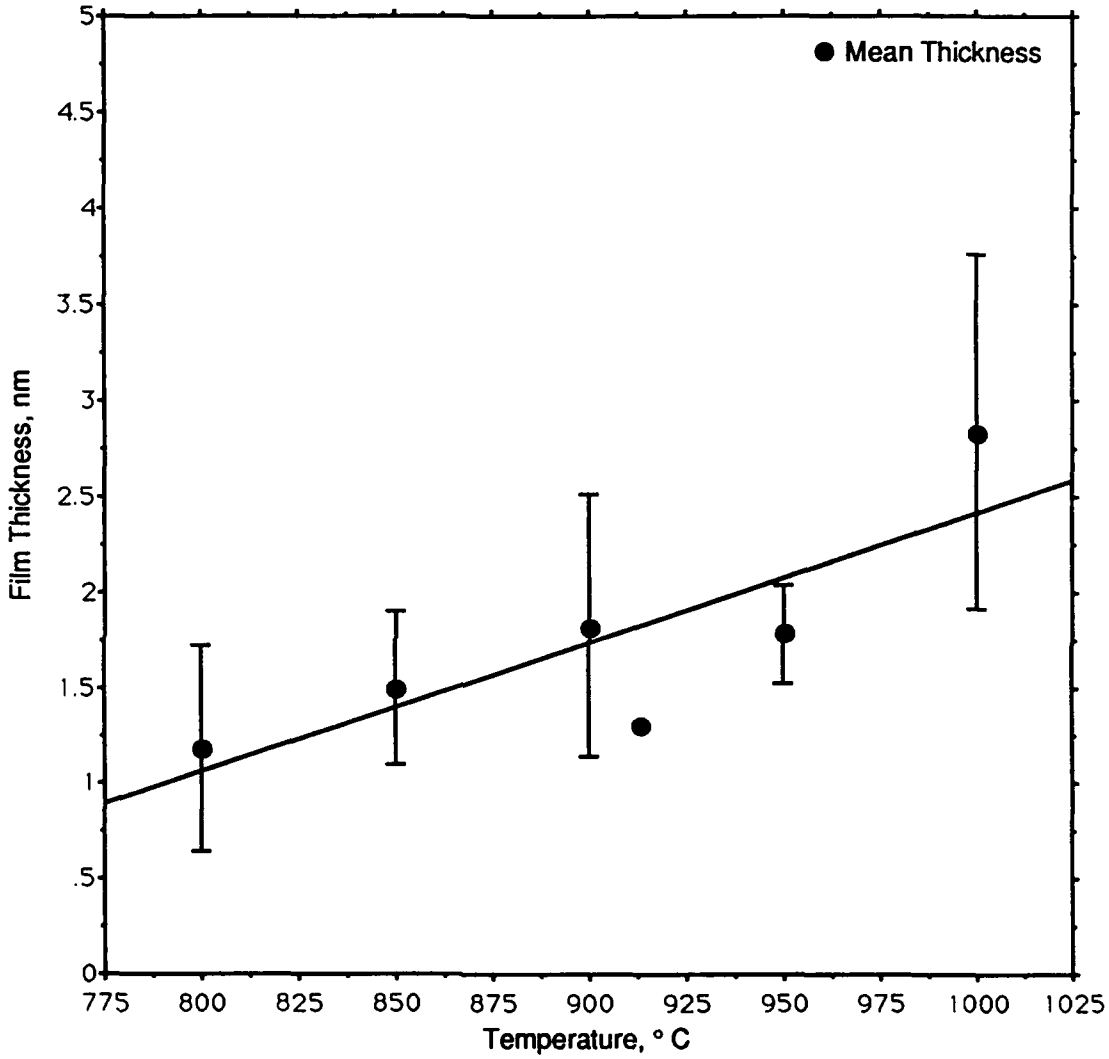


Figure 3. Limiting film thickness for the Region I reaction.

Appendix A

Hardware Operating Notes for the CERL Automated Ellipsometer

****** Hardware Operating Notes for the CERL Automated Ellipsometer ******

This document is a guide to hardware operating characteristics of the Automated Ellipsometer and vacuum chamber designed and built for the Corrosion and Electrochemistry Laboratory, Department of Materials Science and Engineering, The Johns Hopkins University.

These notes document features of the ellipsometer and vacuum chamber, especially the more subtle points of design and operation.

Contents:

1) Introduction

The CERL automated ellipsometer is a dual-mode instrument capable of high speed or high accuracy measurements.

In the high speed mode, polarization modulation ellipsometry is employed. This radiometric technique utilizes a series of lock-in amplifiers to deconvolute the ellipsometric parameters Δ and Ψ from a 50 kHz modulated light (laser) beam. Data rates up to 1000 sec^{-1} are attainable, permitting study of transient surface phenomena such as gaseous oxidation reactions or double-layer formation in electrochemical systems. The technique falls prey to several instrumental limitations, however. Basic accuracy is limited by the digitization process: 1) the ADC employed has 12-bit resolution. This restricts our measurements to one part in 4096. Over the 360° range of a circle, basic accuracy is thus limited to 0.09°. This is almost an order of magnitude greater than that achievable in conventional null mode. 2) sequential scanning of the input channels can cause measurement errors in a dynamic environment (e.g. a rapidly oxidizing surface). I have sidestepped this limitation by an electronic sleight-of-hand. Points acquired at intervals greater than 50 msec (20 sec^{-1}) utilize averaging of 96 samples. The sampling engine is set to 40 kHz which results in complete sampling in 2.4 msec. The amplifier time constants are set to 10 msec, thus any deflection of the signals are so small as to be lost in the digitization noise. The 96 samples are loaded into computer memory as six channels summed 16 times to provide signal averaging (note that I have chosen 16 samples. This ensures that the digitized

sum is not shifted left more than four bits. With 12 bit data, the total value of the sum does not exceed 16 bits, which would result in fatal mathematical errors.).

Conventional null ellipsometry is used to acquire data of both superior accuracy and precision. Switchover between modes is rapid (though the operator must remember to do several things in the proper order), as a two-zone measurement can be obtained in less than one minute.

The vacuum chamber is removable from the ellipsometer rail. This feature is necessary for primary optical alignment and desirable for experimental flexibility. The chamber allows for a wide range of environments; pressures from 1 atmosphere to mid 10^{-9} torr and temperatures from ambient to the melting point of silicon are achievable with the hot stage. Several spare flanges, piezoelectric leak valve control, and an alignment lug inside the chamber also add to experimental flexibility. The large Santovac-filled diffusion pump provides a pumping speed of approximately 70 torr liter sec^{-1} *at the specimen.*

2) Warning

WARNING! This laser is 5 milliwatts. It will blind you if viewed directly. Use proper lab procedures and good common sense at all times.

3) Description of the Optical Chain

The optical chain consists of:

- polarized 5 mW He-Ne laser (632.8 nm)
- two beam steering mirrors
- polarizer (P)
- Hinds PEM-80 photoelastic modulator (M)
- a gap in the rail for insertion of
  calibration polarizer and calibration λ/4 wave plate in PM mode
  λ/4 wave plate in null mode (Q)
- specimen stage
- two exit apertures
- analyzer (A)
- periscope
- two detectors:
  photodiode for PM mode
  photomultiplier for null mode

4) Directions for Alignment of the Optical Rail

Three levels of alignment are required to get quantitatively accurate performance from the instrument. As with any ellipsometer, extreme care is required during alignment. Several primary or fiducial calibrations must be performed to define the optical path. THESE ALIGNMENTS, ONCE COMPLETED, MUST NOT BE DISTURBED. To do so would destroy the integrity of the system and invalidate any attempts at quantitative precision. Primary alignments include the laser orientation, beam steering and exit aperture alignments. Secondary alignments (measurements) comprise those

which define the plane of incidence (note that the rail holding the optical elements makes the plane of incidence of the ellipsometer horizontal, while the specimen is mounted vertically): polarizer azimuthal error (ΔP), analyzer azimuthal error (ΔA), $\lambda/4$ wave plate azimuthal error (ΔQ), calibration polarizer orientation, and modulator orientation are listed here along with the $\lambda/4$ wave plate defining characteristics of phase shift (Δ_c) and transmittance ratio (T_c). The symbols carried in parentheses are the numerical deviations from ideal which can be input into the McCrackin ellipsometry analysis program (1) to permit use of data read directly from the polarizer and analyzer scales. Tertiary or routine alignments include specimen translation and tilt, angle of incidence, modulator drive amplitude, and amplifier gains. The angle of incidence and amplifier gains are typically selected and set up for an entire study. They only require a "preflight check" to be certain that no one has bumped into them. If large temperature changes (>5° C) occur in the room, modulator drive amplitude may require adjustment due to the effects of thermal expansion on the quartz oscillator. Specimen translation and tilt must be adjusted for each sample, and sometimes during experiments.

I have used the terms primary, secondary, and tertiary as a measure of the degree of impact on ellipsometer performance. In addition, performance of tertiary alignment is routine, secondary is inconvenient, and primary requires a complete stripdown of the bench. For easiest alignment, the secondary measurements should be made on a separate rail before proceeding with the primary and tertiary alignments. The numerical values quoted in the text are those obtained by me. They are meant to be both illustrative and useful as a guide when performing the alignment ritual on the CERL ellipsometer.

a) Primary (Fiducial) Alignments

Position the ellipsometer on a stout bench with feet under the two jackscrews. Level the rail by adjusting the jackscrews. With all components removed, place the rail in the straight-through position. Mount the laser, beam steering mirrors, polarizer, and analyzer/periscope on the rail. Lay the laser so that the white fiducial mark is up. This is the laser's plane of vibration, which must be 90° to the plane of incidence. Secure the polarizer and then the laser head. Paste a piece of cellophane tape or suitably foggy material over the exit end of the periscope to allow inspection of the laser beam. Adjust the path of the laser beam so that it passes through the center of both polarizer and analyzer. The reflected light from the various components can be used as an optical lever to aid alignment. It is acceptable, indeed desirable, to avoid having the reflected laser light reenter the cavity; competitive lasing will result with unpredictable effects on stability.

Mount the two exit apertures on the exit rail, moving the analyzer if necessary. When positioning the analyzer, make sure that there will be adequate swinging room for the exit apertures about the cell exit window. Secure the analyzer. Place the photodiode detector in position and adjust the fine beam steering screws to maximize signal intensity. At this point, seal the beam steering assembly to prevent tampering. Adjust the exit apertures to obtain evenly round umbras on both. Lock them down and seal them. This alignment defines the optical path of the exit beam. It will be used to set specimen translation/tilt in all subsequent experiments. The black tube placed between the apertures lessens the effects of dust and air currents on the beam.

b) Secondary Alignments

These alignments require the use of a good sample stage. I used the stage of the Rudolph ellipsometer of the NIST Metallurgy Division. This five-axis stage permits reliably accurate sample positioning for this work. In addition, the beam chopper/lock-in signal detection system provides superior sensitivity for null measurements.

Set the laser source, polarizer, chopper, stage, analyzer, and photomultiplier on the rail and align straight through. Cross the polarizer and analyzer prisms to get minimum signal intensity. These instruments, when rotated, should track each other within 0.02° over the entire 360° range. For our prisms, the measured primary offset, ε, between P and A+90° is 0.55°±0.01°. Place a clean gold surface on the sample stage. Rotate the stage perpendicular to the incoming light beam. Center the sample horizontally and vertically. Adjust vertical tilt until the reflected light goes back into the source by the optical lever technique discussed in section 3a. Turn the sample parallel to the beam and adjust z-translation until half of the beam is intercepted. Turn perpendicular and readjust the vertical tilt. The accuracy of this last adjustment defines the plane of incidence of the ellipsometer and thus directly affects the accuracy of ΔP, ΔA, and ΔQ. Set the angle of incidence to the principal angle of 74.44° ($n = 0.204$, $\kappa = 16.265$ for Au at $\lambda = 632.8$ nm) and rotate the sample to maximize signal intensity. Set A = 90°, parallel to the plane of the surface. Minimize and record the intensity by adjusting P only. Move A by 0.1° and repeat. If the intensity is increasing, go the other way in 0.1° increments. Plot the P and A readings. Also plot the theoretical line P = A-90°. The intersection of these lines is the true A, A', which for these prisms is 90.34°. Graphically, the minimum P, P', is found to be 359.79°. By definition,

$$P = P' + \Delta P$$

$$0 = 359.79° + \Delta P$$

$$\Delta P = + 0.21°$$

$$A = A' + \Delta A$$

$$90 = 90.34° + \Delta A$$

$$\Delta A = -0.34°$$

These values can be input into the McCrackin program (1) as part of the ALINE instruction. The P and A readings taken from the azimuth scales can be input directly into the program and the ALINE corrections will be automatically applied.

Determination of the $\lambda/4$ wave plate fast azimuth is a simple matter. Set the rail to the straight through position with the sample removed. Cross the polarizer and analyzer with P at 45° and A at 135° (for our prisms, set P = 44.79° and A = 135.34°). Insert the $\lambda/4$ wave plate and adjust its azimuth about 45° to obtain the best null. For our plate, Q = 44.80° to give a true Q, Q', of 45°. LOCK THE COMPENSATOR AT THIS VALUE. The $\lambda/4$ wave plate is properly adjusted both for use in null ellipsometry and as a calibration optic for polarization modulation ellipsometry . When using the McCrackin program, assume

$$\Delta Q = 0°$$

fast axis azimuth = 45°.

The $\lambda/4$ wave plate must be characterized. Using the method of ref. 3, set the rail to the straight through position with P = 45° and A = 135°. Adjust A to minimum intensity, then adjust P. Iterate to find the lowest intensity and record P and A.

Repeat with P = 135° and A = 45°. Input these values into the McCrackin program and execute the CWP (Calculate WavePlate) instruction. For the CERL components, my wave plate measurements resulted in

$$\Delta_c = 89.999°$$

$$T_c = 0.779.$$

Calibration polarizer orientation is best made on our ellipsometer rail. After performing all primary and secondary alignments, mount the calibration polarizer on the rail. Adjust horizontal and vertical translations, then rotate the polarizer about the y-axis (its post mount) and shim the base if x-tilt is required. The object is again to make the optical element normal to the beam. Insert a piece of clear mica between the beam steering mirrors to depolarize the laser beam. Set P = 0° (359.79°) and A = 0° (0.34°) and rotate the calibration polarizer in its mount to achieve minimum signal intensity. Lock it down.

The final alignment is modulator orientation. Set the rail to the straight through position, with P = 0° (359.79°), the modulator mounted on a +45° incline such as a small carpenter's square, and A = 45° (45.34°). Set the modulator drive amplitude to any non-zero value. Measure the signal intensity from the real 2ω lock-in amplifier (see section 5). Shim the modulator as required to minimize the 2ω signal. Also, adjust translation and tilt to center the optic and achieve perpendicularity to the beam.

c) Tertiary Alignments

The proper modulator drive amplitude creates a relative retardance of 138.1° to fulfill the Bessel function criterion of $J_0(A) = 0$. See reference 2 for theoretical

considerations of polarization modulation. To determine the required drive amplitude, set the rail to the straight-through position with P = A = 90° (89.79° and 90.34°, respectively). Record the low pass filtered DC signal intensity with the modulator off and call it I_2. Switch the modulator on and adjust the drive amplitude until the lowest DC signal intensity is found. Call it I_3. The following intensity relationships are used to determine I_1, the desired signal intensity:

$$I_1 = I_2/2$$

$$I_1 = I_3/0.6$$

These derived intensities usually agree within 1%. See reference 3 for the mathematical derivation of these equations.

The angle of incidence is the simplest alignment. Swing the exit arm to the desired angle and lock it down. Note that the locking brake applies a torque to the arm and can move it by several hundredths of a degree; it is the operator's responsibility to compensate for this. I only mention this alignment as a reminder to check it periodically, in case it has been disturbed.

Specimen z-translation and x-, y-tilts are the most frequent alignment. See section 8 for a more complete description of hot stage operation. Regardless of the type of stage, alignment is achieved by manipulating translation and tilt to obtain symmetrical umbras about both exit apertures.

d) Final Polarization Mode Settings

For polarization modulation ellipsometry, set P = 90° (89.79°) and A = 315° (315.34°).

5) Description of the Electronics

The polarization modulation ellipsometer is driven by a free-running oscillator, the photoelastic modulator. The optical element resonates at approximately 50 kHz. The drive circuitry outputs two signals, f and 2f, which are used as phase references for the signal processing circuitry. Light detection and signal deconvolution are performed parasitically, meaning that nothing that the operator does to the signal detection chain has an effect on the ellipsometer itself.

The signal chain consists of a MRD555 photodiode, preamplifier with gain and offset corrections, a card rack containing the signal processing circuits, the AT&T 6300 personal computer, and diagnostic oscilloscope and voltmeter.

The photodiode is reverse-biased so that leakage current is proportional to incident light intensity. A 1 kΩ foot resistor limits the current and provides a voltage drop to ground that can be detected by an AD521KD instrumentation amplifier. Gain is switchable at x1 and x10 and a 10-turn potentiometer permits voltage offset adjustment. The output is directed to the signal input on the card cage and to the oscilloscope.

The card cage houses ten circuit boards. Slot 1 contains a digital time base and digital distribution network designed for a modulator not used here, hence this card is not used. Slot 2 contains analog distribution amplifiers and the low pass filter which provides an average DC signal intensity used to normalize the frequency-derived signals described below.

Slots 3 and 5 contain the lock-in amplifiers for the real and imaginary parts of the ω signal, respectively. They are fed intensity signals from the distribution amplifiers on card 2, and reference signals of frequency f from card 4, the ω phase control card. This card is functional and can be controlled from the front panel.

Slots 6 and 8 contain the lock-in amplifiers for the real and imaginary parts of the 2ω signal, respectively. They are fed intensity signals from the distribution amplifiers on card 2 and reference signals of frequency 2f from a small card attached directly to the card cage bus. The 2ω phase control card in slot 7 is not used because the upper frequency limit on the phase control card is 50 kHz, whereas the 2f input is 100 kHz. The imaginary reference component is derived from the small circuit board which contains a 100 kHz quadrature generator. The front panel phase controls are not active for the 2ω signals.

Slot 9 contains the interface card to the computer. The low pass filtered DC and external input signals from card 2, real ω from card 3, imaginary ω from card 5, real 2ω from card 6, and imaginary 2ω from card 8 are transmitted to card 9 across the card cage bus. These signals are relayed down to the analog input multiplexer on the multifunction DASH-16 board in the AT&T 6300 personal computer. Also, front panel gain selection switch positions are sensed by card 9, multiplexed, and transmitted to the DASH-16 via digital I/O lines. This allows the computer to sense the amplifier gains and automatically scale the signal intensities. More information on this process may be found in the file PROGRAM.DOC.

6) Description of the Vacuum System

Growth of device quality silicon oxides demands careful specimen cleaning and control of the ambient environment. The ultrahigh vacuum growth chamber developed for this research has tried to meet these requirements by addressing the following criteria:

- Ultrahigh vacuum (< 10^{-8} torr) for flash anneal
- High pumping speed to minimize time required for flash anneal
- Nude ion gauge in close proximity (5 cm downstream) to specimen
- Good swirl pattern set up during oxygen introduction
- Good vibration damping
- Minimization of window distortion under changes in pressure
- Good mechanical stability for specimen support, even during temperature changes

The specimen chamber consists of a 304 stainless steel pipe spool 4" i. d. x 10.62" in length. The long ends are 6" Del Seal® rotatable flanges. Seven 2 3/4" Del Seal flanges are arrayed about the chamber. Two flanges provide the laser entrance and exit ports; these are rotatable flanges set at an angle of 70° and centered on a target point located 5.62" from the rear end of the chamber. One flange is directly above the sample target point; currently a viewport, it can be used to access specimens in other experiments such as cleavage *in vacuo*. The other four flanges are arrayed on the top side of the chamber 30° off vertical, two behind and two in front of the target point. A piezoelectric leak valve and a spare flange are located behind the target point. One flange ahead of the target holds a thermistor vacuum gauge and a nude ion gauge. The other holds an up-to-air valve and a viewport used for pyrometry readings. Inside the

chamber is a lug aligned with the rear face in the plane of the target point. This lug can be used as an aid in target point alignment or to provide mechanical stability for experiments such as cleavage *in vacuo*. The chamber is mounted on a 1/4-20 thread welded externally below the target point. A large ball-and-socket joint was fabricated to sit on the ellipsometer rail. The curvature of the joint is a 3 1/2" radius so that tilt of the entire chamber would be possible while maintaining reasonable specimen eucentricity.

The windows are 3/8" optically flat and annealed fused quartz, mounted on 1 1/2" of 1/8" thick fused quartz tubing. This tubing meets a graded glass seal to 7052 glass tubing which is, in turn, Kovar-mated to a 304 stainless steel bellows and Del Seal flange. These windows are fully bakeable to 400° C. A brass frame surrounds the bellows to provide three point adjustment for window tilt. No measurable ellipticity was found in these windows, nor any measurable strain birefringence under vacuum.

The chamber is mated to a 4" bellows which is hinged to limit movement to the vertical plane only. This permits movement of the vacuum system and some adjustment of vertical tilt of the growth chamber while preventing lateral tilt and collapse of the bellows under the pressure differential. Additionally, a 10-32 rod and lug is located atop the bellows. BEFORE REMOVING THE CHAMBER FROM THE RAIL, secure the bellows with the screw and nut provided! The chamber is cantilevered off of the pumping station, and will swing down on the bellows hinge, destroying the bellows, windows, gas feed plumbing, and quite probably cause the pumping station to fall over and crush the operator.

Downstream from the bellows is a 10.62" tee containing a spare flange. This is to accommodate additional pumping equipment, such as ion or titanium pumps. The total system volume to this point is approximately 5.8 liters.

Next is a modified 4" i. d. bellows-sealed gate valve used to isolate the chamber from the pumping stack beyond. This valve is manually operated, so that throttling can be achieved if necessary. A small Nupro® leak valve has been inserted into the entry and exit flanges to parallel the main valve. This valve provides fine control of leak rates. I use it when pumping down the chamber; opening the main valve causes the chamber to depressurize rapidly and shocks the windows. This valve will be referred to below (section 8) as the "green valve" (by the color of the knob).

Beyond the gate valve is a 10.62" pipe spool and a 90° vacuum elbow (which helps minimize backstreaming from the pump). Below the elbow is an adaptor to a 4" ASA flange.

The ASA flange is O-ring sealed. This flange is supported by an aluminum shelf, with the pumping stack hanging beneath. The flange is integral with a liquid nitrogen cold trap.

An expanded-mouth 6" i. d. three stage oil diffusion pump with Mexican cold cap hangs from the trap. This pump is capable of 1500 torr liters sec^{-1}. It is charged it with 125 ml of Santovac-5 pumping fluid.

The entire pumping stack is mounted on a portable frame so that the growth chamber can be removed from the ellipsometer and wheeled aside.

7) Description of the Kinematic Hot Stage

The hot stage used in the gaseous oxidation work represents three years of development.

The stage needed to meet several criteria:

- UHV compatibility
- accommodation of specimen thermal expansion/contraction while maintaining the specimen in plane
- ability to withstand extreme temperatures for long times
- low outgassing rate
- electrical isolation
- minimal vibration
- ability to translate and tilt

The base of the stage is a 6" Del Seal flange, which mounts to the rotatable flange at the rear of the specimen chamber. This special order flange contains four Mini-Conflat® flanges arrayed in a square. Three flanges accommodate bellows-sealed linear motion feedthroughs for specimen tilt and translation. The fourth holds a 9 pin instrument feedthrough. Six of the pins are used:

- two pins carry high current DC for specimen heating
- two pins carry AC current for stage bakeout
- two pins connect to a chromel-alumel thermocouple mounted within the stage to monitor bakeout temperature

A 304 stainless steel platen rests on the ends of the linear feedthroughs. The feedthroughs terminate in ball noses which engage the platen in a flat/cone/vee arrangement known as a kinematic mount. This permits simultaneous movement of all three feedthroughs to achieve z-translation of the platen. With the upper left translator fixed, movement of the right translator only results in platen tilt about the vertical axis. Movement of the lower translator only results in tilt about the horizontal axis.

A bakeout heater is mounted on the back of the platen. It is a piece of nichrome wire wound into an element, mounted on Vycor standoffs, and insulated from the stage by mica sheets. A bakeout circuit uses a filament transformer to isolate the heater from line voltage in case of electrical leakage. The transformer is plugged into a lamp timer so that bakeout may be automatically performed at times when the operator is not present. Temperatures of 150° C are attained in one hour of baking. This is adequate at 10^{-8} torr.

The front face of the platen carries the specimen mounts. One side of the specimen is mechanically fixed and electrically isolated while the other accommodates the thermal expansion and is electrically grounded. Two mica-glass (machinable ceramic) end blocks flank the platen to constrain the specimen loading train. The electrically isolated side of the loading train consists of an Inconel 600 contact held by two mica-glass knife edges. The positive lead from the specimen heat power supply is connected to this contact. The contact rests against one of the mica-glass end blocks. After repeated use, the Inconel gets pitted and contaminated with silicon; the contacts can be machined down and regrooved, and a shim inserted into the mica-glass end block to maintain loading spring pressure.

The other side of the loading train consists of a 304 stainless steel carrier slightly (0.002") larger than the Inconel contact. The contact is free to slide in the carrier but cannot slip out of plane. The thermal expansion coefficient of 304 is slightly larger than that of Inconel, hence as the temperature increases the contact will not bind in the carrier. A 304 stainless steel leaf spring (made from shim stock) resides between the contact and the mica-glass end block. The contact on this side of the specimen must be connected to the ground side of the power supply to minimize the chance of a high current ground loop through the instrument chassis.

There is a cutout in the platen behind the location of the specimen. The pyrometric measurements rely on an assumption of non-blackbody conditions; this slot is to allow radiant energy to escape from the sample without being reflected back into it.

8) Specimen Exchange Instructions

The vacuum system is backfilled to 1 atm with water-pumped nitrogen when "cold". Specimen exchange requires removal of the stage. Loosen two opposing bolts at the 3 o'clock and 9 o'clock positions form the specimen stage flange. Remove the washers and loosely reattach the nuts; these bolts serve as a safety when the stage pops free. Remove the remaining 14 bolts completely. Support the stage by cradling it in the right hand with thumb and forefinger on the safety bolts. Place a small screwdriver into the helium leak check groove at 12 o'clock and pry the flanges gently apart. The stage will pop loose as the metal gasket parts from the chamber. Spin the nuts from the safety bolts and remove the stage, using the bolts for support and guidance while accommodating to the weight of the stage. Place the stage platen-up on a large plastic

beaker or other suitable support. Hands should be washed in order to avoid contaminating the UHV surfaces, especially with the black anti-seize compound found on the bolts.

Place the aluminum support jig under the specimen. This jig is 0.215" thick to provide support for specimen mounting. Draw back the spring-loaded contact and remove the specimen with a pair of wafer tweezers. Replace the contact blocks with a freshly prepared set (there are 6 identical blocks). Place a new specimen on the alignment jig, withdraw the spring-loaded contact until the sample falls into place, release, then withdraw the support jig. Slide the sample gently in the grips to assure firm, continuous contact. Measure specimen resistance at the feedthrough; it should be less than 100 ohms. Replace the copper gasket with a fresh one, handling it by the outside edge only. Gently blow contamination from the specimen using compressed gas.

Remount the stage, securing it firmly with the two safety bolts. Replace all bolts. Tighten opposing pairs of bolts in a random fashion to avoid chasing the metal gasket around the flange. Torque in three stages to 45 lbs.-ft. After completion of the torquing sequence, tilt/translation adjustments may be made. Connect DC and AC power leads. Start the rotary pump if it is not already running, then open the green valve for two minutes. Turn on the cooling water for the diffusion pump and plug in the pump. Open the gate valve and close the green valve. After 15 minutes, the nude gauge can be energized. From a cold start, it takes about 3 days to reach operating conditions. From a warm start, turnaround is usually overnight, though the cycle can be pushed to 3 hours.

Note that pumpdown requires that the diffusion pump be shut down. There is no provision for direct rotary pumping of the sample chamber. This design reduces

pumping system complexity, saves money, and ensures a cleaner environment at the specimen.

9) Execution of Experiment

Making Polarization Modulation Ellipsometric Measurements

Ensure that P = 90° (89.79°) and A = 315° (315.34°), the modulator is on and the signal is maximized on the oscilloscope by tilt/translation of the specimen. A voltmeter is connected to the low pass filtered DC signal monitor, the output of which should read 1.8-2.0 volts. Consult SOFTWARE.DOC for instructions on operating the polarization modulation ellipsometer.

Making Null Measurements

Shut off the photoelastic modulator. Place the $\lambda/4$ wave plate (Q) on the rail. Rotate P and A to ballpark values e. g. for native SiO_2 P = 48° and A = 10.5°. Flip the periscope so that the photomultiplier intercepts the laser light. Obtain your P and A readings. These are zone 2 readings. Reverse the $\lambda/4$ wave plate on the rail (do not change the azimuth–turn the plate around) to get the zone 4 measurements. Using two-zone readings, one can execute the McCrackin program CAT (Compute Angle of Tilt) instruction to correct for the specimen shifting in the grips.

Reset the periscope to the photodiode. Set P = 90° (89.79°) and A = 315° (315.34°). Turn the modulator on. After <5 seconds it will return to its previous value with no measurable shifts in retardance. Remove Q from the rail.

Example: Gaseous Oxidation

After native oxide measurements have been taken by both means above, one can perform a flash anneal and grow an oxide.

Set the pyrometer to 1125° C. This corresponds to a true sample temperature of 1204° C. Turn on the specimen power supply with the current limiter at minimum and the voltage limiter at maximum. Observe the specimen through the pyrometer. Increase the current to the sample. Initially, the voltage will increase as the sample has some resistance (dependent on dopant level). At about 580° C thermal electron-hole pair generation takes over and the specimen becomes a short circuit. The output voltage will fall and current limiting comes into play. The specimen will begin to glow. Temperature is now a function of applied current. Adjust the current until the sample brightness matches the pyrometer filament. This whole exercise should take about two seconds. Monitor the chamber pressure; it will increase by at least an order of magnitude, then start to fall after 6-8 seconds. 10-15 seconds is adequate to flash the specimen. Shut down the current. The specimen tilt will require readjustment. Take readings to establish the cleanliness of the surface.

Ready the PME to take rapid (e. g. 100 msec) data. Set the pyrometer to the experiment temperature. Bring the specimen up to temperature. Tweak tilt alignment. TURN OFF THE NUDE ION GAUGE. Close the gate valve and make sure the green valve is also closed. Hit <return> to start data acquisition, wait about 20 seconds to establish a baseline, then open the up-to-air valve to admit ultrahigh purity oxygen. Tweak tilt if necessary with quick, deliberate movements. While adjusting tilt, junk

data is accumulated. If the junk is confined to short spikes in the data, it can be more readily identified as such and filtered out later.

After about one minute, stop acquisition and take null measurements. Restart the PME at the desired data rate (e. g. one point per minute). Check the temperature periodically. Shut down the diffusion pump so that it can cool adequately.

10) Pyrometric Measurement

Both the real and imaginary parts of the complex index of refraction of silicon are functions of temperature, thus any useful conclusions from ellipsometric observations require accurate temperature determination. I chose optical pyrometry (650 nm wavelength) with careful attention to two correction factors. First, *direct observation* of a sample in a vacuum system is not possible. At least one window must be present to maintain vacuum, and attenuation of the light by that window will result in an underestimation of sample temperature. The construction of the CERL vacuum chamber required the use of one window and two mirrors which further increases the potential for measurement error. Experimental attenuation measurements were carried out by loading a silicon sample in the chamber with the ellipsometer entrance window removed. The pyrometer was sighted through the opening to permit direct observation of the sample. The secondary mirror was realigned so that the pyrometer could swing to sight the sample via the mirrors. The specimen was heated, then measured and recorded by each optical path. Several different temperatures were measured, providing data from which attenuation corrections ("apparent" temperatures) were derived by Wien law calculations (4). Second, emissivity corrections were applied for

the non-blackbody conditions in our cell. Van der Meulen and Hien (5) published useful values for n and k of silicon at elevated temperatures. Non-linear least squares fits to their data were incorporated into a small computer program which could provide interpolated n and k for given apparent temperatures. Emissivity values were then calculated from these optical constants.

Wien law calculations based on this emissivity produced a "true" surface temperature, which was then fed back into the program to refine selection of the optical parameters. Convergence to true surface temperature occurred after three program iterations.

11) Glossary of Terms

A — Analyzer. An optical component that linearly polarizes light; a polarizer. Specifically, the linear polarizer found at the "downstream" end of the optical chain, used to cross-polarize the emergent light beam from the specimen and thereby attain null intensity. It is easier to measure a minimum intensity than a maximum.

ΔA — The difference between the azimuthal orientation of the analyzer's optical element and the scale reading for that element. Rather than adjust the scale to read the exact value of the analyzer, one measures this quantity and applies it to the scale readings obtained subsequently. This correction is performed by the McCrackin software. The mathematical definition is given in section 4b, Secondary Alignments.

ADC — Analog to Digital Converter.

CERL The Corrosion and Electrochemistry Research Laboratory at The Johns Hopkins University. The lab is a working group consisting of faculty and students from the Department of Materials Science and Engineering and the Department of Chemical Engineering who share common research interests in corrosion science and engineering.

Δ Delta. This is one of the two data obtained from ellipsometric measurement. Δ is the phase difference between two orthogonal components of the probe light beam. These components are r_p, that component parallel to the plane of incidence, and r_s, the component perpendicular to the plane of incidence. The mathematical definition is

$$\Delta = (\beta_p - \beta_s)_{reflected} - (\beta_p - \beta_s)_{incident}$$

where β is the absolute phase of each component of the light.

Δ_c Delta-C. This is a performance parameter of the quarter wave plate, Q. Q is very rarely a true quarter wave plate; Δ_c is the measured retardance of Q. Knowledge of this value is required to obtain exact solutions to the Drude equations (the basis of ellipsometry). It can be input into the McCrackin program.

f The fundamental frequency applied to the photoelastic modulator. This signal is sent to the lock-in amplifiers as a reference for detection of the ω signal from the photodiode.

2f The second harmonic of the frequency applied to the photoelastic modulator. This signal is sent to the lock-in amplifiers as a reference for detection of the 2ω signal from the photodiode.

k The imaginary, or absorbing, part of the index of refraction of a material.

κ A different way of stating the imaginary part of the refractive index of refraction. κ is related to k by the relation

$$\kappa = k/n$$

M Modulator orientation. This is the angle between the long axis of the photoelastic modulator and the plane of incidence.

n The real part of the index of refraction of a material.

P Polarizer. An optical component that linearly polarizes light. Specifically, the linear polarizer found at the "upstream' end of the optical chain, used in conjunction with the $\lambda/4$ wave plate to create light with a phase shift of $-\Delta$ for null ellipsometry. The Polarizer, as with other linear optical components on the rail, is oriented to zero when the plane of vibration of the light is coincident with the plane of incidence of the ellipsometer. For our rail, this means horizontally.

ΔP The difference between the azimuthal orientation of the polarizer's optical element and the scale reading for that element. Rather than adjust the scale to read the exact value of the polarizer, one measures this quantity and applies it to the scale readings obtained subsequently. This correction is

performed by the McCrackin software. The mathematical definition is given in section 4b, Secondary Alignments.

Plane of Incidence. The plane which contains both the impinging and emerging light beams reflected from a specimen. This plane defines the absolute azimuth of the ellipsometer.

PME Polarization Modulation Ellipsometry. This technique phase modulates a light beam (rather than fixing the phase difference at 90°) to obtain Δ and Ψ. Much higher data rates are achievable over null ellipsometry, but with a sacrifice of both accuracy and precision.

Ψ Psi. This is one of the two data obtained from ellipsometric measurement. Ψ is the amplitude ratio between two orthogonal components of the probe light beam. These components are r_p, that component parallel to the plane of incidence, and r_s, the component perpendicular to the plane of incidence. The mathematical definition is

$$\tan\Psi = \frac{|r_p|}{|r_s|}.$$

Q Quarter-wave plate. Often referred to in the text as a $\lambda/4$ wave plate or compensator. An optically active substance which retards light polarized in one direction (such light is called the "ordinary ray", which lies along a "slow axis") relative to light polarized in another direction (called the "extraordinary ray", which lies along a "fast axis"), cut to the proper thickness d so that the relative retardance is 1/4 of a wavelength of light.

Normally, such plates only meet their performance criteria at one specified wavelength, λ, as given by the equation

$$\delta = (2\pi/\lambda)d(n_O - n_E)$$

The quarter wave plate is oriented to zero when the plane containing the fast axis lies coincident with the plane of incidence of the ellipsometer. For our rail, this means horizontally.

ΔQ The difference between the azimuthal orientation of the quarter wave plate's optical element and the scale reading for that element. Rather than adjust the scale to read the exact value of the quarter wave plate, one measures this quantity and applies it to the scale readings obtained subsequently. This correction is performed by the McCrackin software. The mathematical definition is given in section 4b, Secondary Alignments.

T_c This is a performance parameter of the quarter wave plate, Q. T_c is the measured transmittance of Q. Quarter wave plates frequently attenuate the ordinary and extraordinary rays to different extents. Knowledge of this value is required to obtain exact solutions to the Drude equations (the basis of ellipsometry). It can be input into the McCrackin program.

ω The intensity component of frequency ω measured by the photodiode in polarization modulation mode. This signal is related to the ellipsometric parameter Δ. See reference (5) for the mathematical details.

2ω The intensity component of frequency 2ω measured by the photodiode in polarization modulation mode. This signal is related to the ellipsometric parameter Ψ. See reference (5) for the mathematical details.

12) References

1. F. L. McCrackin, NBS Technical Note 479, U. S. Government Printing Office, Washington, DC (1969).

2. S. N. Jasperson and S. E. Schnatterly, *Rev. Sci. Instrum.*, **40**, 761(1969).

3. P. J. Hyde, Polarization Modulation Ellipsometer as Installed at Los Alamos National Laboratory, Los Alamos National Laboratory (1983).

4. W. P. Wood and J. M. Cork, Pyrometry. McGraw-Hill, New York (1927).

5. Y. J. van der Meulen and N. C. Hien, *J. Opt. Soc. Am.*, **64**, 804 (1974).

The following references are not specifically cited, but are recommended as good general references on ellipsometry:

6. F. L. McCrackin, E. Passaglia, R. R. Stromberg, and H. L. Steinberg, *J. Res. NBS A*, **67A**, 363 (1963).

7. D. E. Aspnes and A. A. Studna, *Appl. Opt.*, **10**, 1024 (1971).

8. W. A. Shurcliffe, Polarized Light, Harvard University Press, Cambridge, MA (1962).

9. F. A. Jenkins and H. E. White, Fundamentals of Optics. McGraw-Hill, New York (1976).

Appendix B

Software Operating Notes for the CERL Automated Ellipsometer

****** Software Operating Notes for the CERL Automated Ellipsometer ******

This document is a guide to software operating characteristics of the Automated Ellipsometer designed and built for the Corrosion and Electrochemistry Laboratory, Department of Materials Science and Engineering, The Johns Hopkins University.

These notes are intended to serve as a user guide. Details of program design are found in the file PROGRAM.DOC. Hardware notes can be found in the file HARDWARE.DOC.

Contents:

1) Introduction

The software written to control the CERL Automated Ellipsometer is meant to be user-friendly and simple to operate. The program choices (flexibility) are deliberately limited to minimize operator confusion. Certain operations such as file naming are handled automatically to provide consistency and robustness; my philosophy has been to let the computer log run times, calibration constants, et cetera, as a backup to a tired and overworked operator.

Several options and features are available:

- two operating modes
- automatic signal averaging on slower runs
- automatic sensing of lock-in amplifier settings
- automatic file naming in day/month/year/run format e. g. 10MAR58.001
- automatic logging of start/end date and time
- automatic query to operator for a comment string
- external input available e. g. to log potentiostat voltage

2) Loading and Execution

As presently configured, an AT&T 6300 computer controls the instrument. The program ELLIPS.EXE must be loaded to capture data.

The current AUTOEXEC.BAT delivers the operator to the directory TURBOC. This is the development environment. It is important to note that graphics drivers are located here, defined by the extension .BGI . When the program starts to execute, it automatically detects the type of display and looks into the directory where it resides for the appropriate graphics driver. *If you plan to move the program ELLIPS.EXE to another directory, you must move the appropriate graphics driver with it.*

To load: simply type the command ELLIPS<return>. Execution begins immediately.

3) INITIALIZE Screen

Upon startup, the program must initialize. The title "INITIALIZE" appears in the upper right-hand corner of the screen, a feature implemented so that the operator doesn't get lost in the hierarchical menus. Several messages appear, which should be of little interest if things are working properly. First, the exit directory is established so that on program termination the user is back in TURBOC:

```
Exit directory established.
```

A file storage path is established, currently the same directory as the exit:

```
File storage path established.
```

Memory for data is then allocated. Up to 384 kilobytes of memory are set aside. If there is a problem, you can't run, so the program advises you:

```
Unable to allocate memory.  Sorry, but you HAVE to
quit.

Strike any key to exit....
```

Otherwise, the memory is allocated:

```
6 segments of memory allocated for data.  This is
enough to hold 32766 points, or about 32 seconds at the
maximum rate.
```

The DASH-16 data acquisition board is initialized:

```
A/D board successfully initialized.
```

At any point in the program, if the A/D board responds to commands with a nonzero error message, the operator is notified and the program terminates. Unfortunately, all data will be lost.

The DASH-16 board must be told which channels to scan:

```
A/D scan limits successfully loaded.
```

Hardware interrupts are enabled. No message is given. The smoothing boxcar is flushed, and again no message is given. Upon the following message, the program has successfully initialized:

```
Three... Two... One... Blast off!!!!!
```

INITIALIZE only occurs at the start of program execution; the MAIN screen appears next.

4) MAIN Screen

The message MAIN appears in the upper right-hand corner of the screen. The screen appears as follows:

```
Welcome to the CERL Automated Ellipsometry System.
Your options are:
```

```
1)  Acquire data
2)  Review stored data
3)  Change storage path
4)  Change display options (not active)
5)  Calibrate external input (not active)
6)  Exit to DOS

The current storage path is C:\TURBOC

The amount of memory left is 65112 bytes.

Enter your choice:
```

Option (1) will take you to the ACQUIRE screen.

Option (2) will take you to the LOADFILE screen.

Option (3) will take you to the CHANGE PATH screen.

Options (4) and (5) are not implemented at this time, as my gaseous oxidation work did not require them. Future students may wish to modify this work. I hope that my thought patterns are sane enough so that they find success.

"Exit to DOS" requires confirmation:

```
Exit to DOS (Y/N)?
```

The operator must enter "Y" or "y" to exit, otherwise the MAIN screen is redrawn.

5) ACQUIRE Screen

ACQUIRE appears in the upper right-hand corner.

```
There are two modes of data acquisition:

 Manual:    each point is taken in response to <return>
 Automatic: data are collected at user-selected
 intervals

Yours options are:
```

```
1)  Manual Acquisition
2)  Automatic Acquisition
3)  Exit to MAIN level

Enter your choice:
```

The manual mode is useful for studies that require operator intervention, such as specimen manipulation or manual adjustments to applied potential. Automatic is the more commonly used mode, permitting unattended data logging. I have used the system for experiments of up to forty hours duration without problems.

6) LOADFILE Screen

LOADFILE appears in the upper right-hand corner.

```
Enter input filename (path spec if required):
```

Suppose you wanted to review the file 10MAR58 001, taken on my birthday and just a bit after Sputnik. Type in

```
10MAR58.001<return> or
A:\10MAR58.001<return>
or whatever is required
```

The target directory is searched. If the file is not found, you will see

```
This file does not exist!
```

This is one of the few non-fatal errors in the program. You will be returned to the previous screen. If the file is located, you will see this instead:

```
Loading file 10MAR58.001...
```

Once the data is loaded, the graphics display is activated. Upon hitting <return>, the display is cleared and the previous screen is redrawn.

7) CHANGE PATH Screen

CHANGE PATH appears in the upper right-hand corner.

```
The current storage path is C:\TURBOC

Enter new path:
```

Enter a new storage path specification, such as "A:<return>". Striking the <return> key *only* will abort the routine and return you to the previous screen.

```
Type <return> to accept, <Esc> to cancel, any other key
to try again...
```

The program accepts your input and returns you to the previous screen.

8) MANUAL Screen

MANUAL appears in the upper right-hand corner.

```
Manual Acquisition Mode

 This mode permits manually controlled data acquisition
 in response to <return>.

The external channel is currently disabled.
Smoothing is currently inactive.

Your options are:

  1)  Enable/disable external input
  2)  Enable/disable 21 point boxcar smoothing
  3)  Calibrate ADCs   <== You must do this before
      selecting  (4)
  4)  Run a scan
  5)  Exit to ACQUIRE menu

Enter your choice:
```

Option (5) returns you to the previous screen.

Options (1) and (2) toggle on and off. The display will be redrawn in response. If the external input is not being used, leave it off. The data is still acquired, but it is not stored to disk, saving a great deal of space.

Option (3) is required to provide calibration values to the data reduction algorithms. If the constants have been defined, the bold type will not appear. I try to calibrate at least once per experiment. Choosing (3) will take you to the CALIBRATE screen (see section 9).

Option (4) will take you to the next MANUAL screen. MANUAL will appear in the upper right-hand corner:

```
Manual acquisition mode

Make your final specimen alignments, set amplifier
gains as required, then block the beam for a dark
current measurement.  Strike any key when ready...
```

The program pauses while the operator makes last-second adjustments to the experiment. The darkcurrent reading is needed to subtract out stray light readings and amplifier offsets.

```
Hit <return> to take a point, any other key to abort...
```

Each time the return key is pressed, a set of samples is acquired and the following message appears:

```
Point n taken
```

Striking any other key halts acquisition and invokes the file storage routine:

```
Enter any information to be logged with these data (64
characters max):
```

Input anything you wish, such as

```
Sample 123-4, T=1000, run looks good<return>
```

The program, on error, responds with an appropriate error message. This is a non-fatal error, but the data of this particular run is unavoidably dumped. Otherwise, the program responds

```
Saving file 10MAR58.002
57 points collected in 88 seconds
Another run?
```

Affirmative reply keeps you in this screen. Any other reply takes you back to the first MANUAL screen.

9) CALIBRATE Screen

CALIBRATE appears in the upper right-hand corner. This screen is interactive.

```
Block the beam for a stray light measurement.  Strike
any key when ready...
```

Insert something opaque into the light path. I use a small piece of cardboard, which fits on the rail between the beam steering mirrors. Strike a key.

```
Acquiring dark current data...
```

After a few seconds, the computer beeps, and the above message is replaced by:

```
Dark current = 10452

Insert the calibration linear polarizer between
modulator and sample, set amplifier gains as required,
then strike any key when ready...
```

Follow the instructions.

```
Acquiring '2 Omega' data....
```

In the unlikely event that the DC signal is exactly zero (which results in a fatal mathematical error), the computer displays the following and returns to the previous screen without a valid calibration:

```
Error:  DC level = 0, which is bad.
Strike any key to continue...
```

Otherwise, the following appears:

```
DC = 182934    Re2Omega = 300229    Im2Omega = 1955288
2 Omega calibration vector magnitude = -3.247119

Insert the quarter wave plate between the modulator and
calibration polarizer, reset amplifier gains if
necessary, then strike any key when ready...
```

Similar stuff happens, resulting in the display

```
Acquiring 'Omega" data....

DC = 3211    ReOmega = 342096    ImOmega = 1907
Omega calibration vector magnitude = -2.868113

The calibration is complete.  Remove both calibration
elements from the beam.  Strike any key to continue...
```

The operator is successfully returned to the previous screen.

10) TIMED ACQUISITION Screen

TIMED ACQUISITION appears in the upper right-hand corner.

```
Timed acquisition mode

 This mode permits continuous timed acquisition of
 data.  The acquisition rate may be adjusted in one
 millisecond intervals ranging from one millisecond to
 3,600,000 milliseconds (1 hour).

The external input is currently disabled.
Smoothing is currently inactive.
```

```
The current time interval is 1 millisecond.
Maximum acquisition time is about 32.77 seconds.

Your options are:

  1)  Change acquisition time
  2)  Enable/disable external input
  3)  Enable/disable 21 point boxcar smoothing
  4)  Calibrate ADCs   <== You must do this before
      selecting  (5)
  5)  Run a scan
  6)  Exit to ACQUIRE menu

Enter your choice:
```

Option (6) returns you to the previous screen.

Option (1) lets the operator change the acquisition time to anything between 1 and 3,600,000 milliseconds, and must be entered in this format. See CHANGE TIME INTERVAL in section 13.

Options (2) and (3) toggle on and off. The display will be redrawn in response. If the external input is not being used, leave it off. The data is still acquired, but it is not stored to disk, saving a great deal of space.

Option (4) is required to provide calibration values to the data reduction algorithms. If the constants have been defined, the bold type will not appear. I try to calibrate at least once per experiment. Choosing (4) will take you to the CALIBRATE screen (see section 9).

Option (5) will take you to the next TIMED ACQUISITION screen. TIMED ACQUISITION will appear in the upper right-hand corner:

```
Timed acquisition mode

Make your final specimen alignments, set amplifier
gains as required, then block the beam for a dark
current measurement.  Strike any key when ready...
```

The program pauses while the operator makes last-second adjustments to the experiment. The darkcurrent reading is needed to subtract out stray light readings and amplifier offsets.

```
Hit <return> to start, any other key to abort...
```

There are two possible screens which may now appear. If the scan interval is greater that 56 milliseconds, the computer will have sufficient time to acquire, average, reduce, and plot the data. For shorter intervals (faster data rates), direct memory access is invoked, no signal averaging takes place, and the data are not displayed immediately. The former case is described in section 11 and the latter in 12.

11) GETTDATA Screen

Slower data can be graphed in real time. The screen enters graphics mode and data is displayed as it comes in. Horizontal autoscaling is invoked as necessary. To end the run, strike any key. There may be a delay, up to the time required to acquire the next point, before the file storage routine is called:

```
Enter any information to be logged with these data (64
characters max):
```

Input anything you wish, such as

```
Sample 123-4, T=1000, run looks good<return>
```

The program, on error, responds with an appropriate error message. This is a non-fatal error, but the data of this particular run is unavoidably dumped. Otherwise, the program responds

```
Saving file 10MAR58.002
176 points collected in 88 seconds
Another run?
```

Affirmative reply keeps you in this screen. Any other reply takes you back to the first TIMED ACQUISITION screen.

12) GETFDATA Screen

Faster data cannot be displayed. The previous screen is not erased, but the following line is appended:

```
nPoints = n
```

where n is the point number. It is continuously updated. Striking any key halts acquisition. There will be a time lag (possibly several minutes) while the data are accessed from memory and reduced. The graphics routines will be invoked and the data will be displayed. The storage routine is called:

```
Enter any information to be logged with these data (64
characters max):
```

Input anything you wish, such as

```
Sample 123-4, T=1000, run looks good<return>
```

The program, on error, responds with an appropriate error message. This is a non-fatal error, but the data of this particular run is unavoidably dumped. Otherwise, the program responds

```
Saving file 10MAR58.002
8846 points collected in 88 seconds
Another run?
```

Affirmative reply keeps you in this screen. Any other reply takes you back to the first TIMED ACQUISITION screen.

13) CHANGE TIME INTERVAL Screen

TIMED ACQUISITION appears in the upper right-hand corner. An interactive screen appears:

```
The current time interval is 1 millisecond.
Enter a new time interval (in milliseconds):
```

Enter a value, such as 30000 (half a minute). An invalid value will prompt the error message below, then beep and return to the previous screen.

```
Time value is out of range!
```

Normally, the following appears:

```
The new time interval will be 30000 milliseconds.
Counter 1 = 2000   Counter 2 = 2500

Is this acceptable?
```

"Y" or "y" will effect the change; anything else will not.

Appendix C

Program Notes for the CERL Automated Ellipsometer

****** Programmer's Notes for the program ELLIPS.EXE ******

This document is a guide to the program ELLIPS.EXE written for the Corrosion and Electrochemistry Laboratory, Department of Materials Science and Engineering, The Johns Hopkins University.

These notes are to clarify and supplement the documentation provided within the source code files.

Contents:

reduce.c

file_m.c

plotdata.c

dast_m.c

7) Known Bugs

1) Using Turbo C 2.0

Turbo C 2.0 is a fairly complete language for the IBM and compatibles world. In addition to the Turbo C library, there is a wealth of graphics capability available to the programmer.

I had to develop this project to take data for my doctorate; the program was not the object of the research, but only one tool of many which I needed to get the numbers. As a consequence, the graphics routines (found in plotdata.c) are fairly primitive. They function simply to give the experimenter visual confirmation that his data are reasonable. Other programs can access the data for reduction and presentation.

2) Use of the Medium Memory Model

Turbo C offers a choice of six memory models which define the pointer constructs used to define locations of both code and data. I have 'chosen' the medium memory model because the MetraByte function call TCM_DASG is compiled in the medium memory model. It turned out to be fairly difficult to get everything to work.

The medium model defines code segments with far pointers, hence the code can exceed the 64 kbyte segment limit, up to 1 Mbyte. Data, on the other hand, is defined by near pointer constructs and is limited to one 64 kbyte segment. This isn't really a problem, because 64 kbytes of static data is a LOT of stuff.

But how is the data stored? I wrote my own memory allocation routines. In the module dast_m.c there are three routines:

```
int far *intptr(int)
float far *floatptr(int)
void allocseg(void)
```

allocseg() sets up high memory to receive the data, using the farmalloc() function found in the Turbo library. intptr() and floatptr() access specific regions of the storage area, based on the argument supplied (the point number). Whenever the data need manipulation, movedata() is invoked to bring the data to low memory and send it back. This is a function in the Turbo library. I initially used memcpy(), but found that it gave unpredictable results when used with the medium memory model.

Study these routines carefully. I'm pretty proud of them, and they can teach a lot about C programming.

3) Compilation and Linking

Since this program depends heavily on the call TCM_DASG() which controls the DASH-16 board, we cannot use the Turbo C integrated development environment for

compilation and linking. The command line equivalent TCC is used instead, followed by the link with TLINK. The whole process is an annoyance. I have tried to ease the pain by creating two files,

makefile.

response.

to automate the process. Unfortunately, you have to go in and out of TC and then MAKE to examine each change.

The Borland folks have helped the hapless programmer by providing the MAKE command. MAKE is equivalent to the <F9> key in the integrated environment. It calls TCC for any source files that supersede their .OBJ files, then links the whole project if necessary.

TO USE TCC: simply type 'make<cr>' from DOS and watch it go. 'make' calls a special batch file, 'makefile.' which contains the compilation instruction:

```
.c.obj:
tcc -c -mm $<
```

which says, in effect, "Compile only (-c), using the med m memory model

(-mm), any .c files listed for this project and give the output files the extension .obj".

The list is provided AFTER this command. The first file listed is ellips.exe, which will be the final name of the project, followed by the .obj files required (backslashes symbolize continuation on the next line). Here's the list:

```
ellips.exe: ellips.obj init_m.obj cal_m.obj acquire.obj manual.obj \
  timedacq.obj file_m.obj reduce.obj plotdata.obj dast_m.obj
```

The last line invokes the turbo linker TLINK. It also needs a special batch file, called 'response.'. The call is:

```
tlink lib\c0m @response
```

which says "Call the Turbo linker and use the medium memory model C stuff, plus the stuff in the file 'response.'".

Let's look into the file 'response.' It consists of four lines. The first line is a list of the .obj files to be called for the link. In our case it is too large to fit on one line, so we use the plus sign as we used the backslash above:

```
ellips init_m cal_m acquire manual timedacq file_m reduce+
plotdata dast_m
```

The next line is the name of the executable file, which for us is

```
ellips   (which will become ellips.exe)
```

Line 3 contains the name of the map file, useful for debugging:

```
ellips   (which will become ellips.map)
```

Line 4 specifies the libraries to be used:

```
lib\fp87 lib\mathm lib\graphics lib\dasg lib\cm \x\c
```

We will call in the floating point library for the 8087, medium memory model math routines, graphics routines, the DASH-16 library, and the C medium memory model library. We then set some options switches: \x tells tlink not to generate the .map file, and \c tells it to treat lower case as significant in symbols. Thus 'Fred' and 'fred' are different to the linker.

For more information, consult Appendix D of the Turbo C v.2 Reference Guide.

4) Global Variable Declarations

Global variables are used rather sparingly in this program. C intentionally tries to make individual routines independent for portability reasons. I have tried to follow this philosophy with partial success. It's hard to make the code truly portable when graphics, memory management, interrupt management, and prototype card (the acquisition board) considerations wed the programmer inextricably to one architecture. Hopefully, this code will be portable enough to compile and execute on any IBM compatible machine.

Booleans:

The header file "gad.h" contains the #define statement

```
#define boolean int
```

which simply says that any variables defined as booleans are actually integers. This little trick is intended to make the programmer visualize these variables as booleans and to clarify their functions.

calvalid — is TRUE when valid calibration values are available to the program. It can only be made TRUE by successful completion of the routine calibrat().

extchan — is TRUE when external channel data is valid. The external channel is always acquired, but if the data is junk there isn't any point in storing it to disk.

smoothing — is TRUE when data smoothing is desired. A 21-point boxcar algorithm is applied to the raw data BEFORE it is reduced, which can help remove noise from the measurement.

DASHintset — is TRUE when the interrupt service routine ServiceMode6() is installed. This boolean is a flag, primarily detected by the control-break handler c_break() or DASH-16 error handler, used to restore the proper interrupt vector so that other computer operations won't get zapped.

KBintset — ditto for the keyboard interrupt service routine ServiceMode7().

Integers:

dasaddr — is the hardware address of the DASH-16 data acquisition board. It is defined in the routine ellips.c from a constant declaration DASADDR and is used in initall() to program the board's status register. DASADDR must agree with the switch settings on the board!

irqlevel — is the interrupt level of the DASH-16. It is defined as above, by the constant IRQLEVEL. Valid data are 2 through 7, which reflect the IRQ (Interrupt Request) lines on the PC bus. When an IRQ is asserted, the 8259A Interrupt Controller arbitrates and flags the 8086 with a software interrupt. irqlevel is used in initall() to program the DASH-16 and in Install_DASH_ISR() and Remove_DASH_ISR() to reprogram the appropriate interrupt vector.

dmalevel — is the DMA (direct memory address) level of the DASH-16. When effecting DMA data transfers, a device must access the 8237 DMA Controller on one of four bus address lines. Level 0 is used by memory refresh, level 2 by the floppy drives, and level 3 by the fixed disk. Hence we use level 1. Note that the value defined in the constant declaration DMALEVEL must agree with the switch setting on the board!

lowlimit — is the lower channel limit loaded into the DASH-16 multiplex control register. Adjust it if you wish (it is defined in the constant declaration LOWCHAN), but the software requires six contiguous channels to operate, in the following order:

i) DC signal level

ii) real component of ω signal

iii) imaginary component of ω signal

iv) real component of 2ω signal

v) imaginary component of 2ω signal

vi) external channel

hilimit — is the upper channel limit.

channels — is simply defined (in initall()) as

hilimit-lowlimit+1

acqmode — defines the current acquisition mode. It is set in routine acquire() and used in storfile() as a flag for data-reading routines. It is stored in the data file.

maxpts — is the maximum number of points we intend to acquire. It is defined from the constant declaration MAXPTS in module ellips.c. The basis for the value 32766 is based on six channels requiring twelve bytes storage per channel, or 5461 points per 64 kbyte segment (two bytes are wasted in each segment). We allocate up to six segments for incoming data, hence 32766 total data. maxpts is modified in allocseg() (in module dast_m.c) as memory dictates.

runnumber — is used to construct the name of the data file for a particular run. It is defined in the routine findfile() in module file_m.c.

nblocks — is the number of memory segments allocated for data storage. It has a value between 0 and 6. This value is determined in routine allocseg() (dast_m.c) and is used in routine initall() (where if it is equal to zero, the user must quit) and the data collection routines.

block[] contains the actual values of the segment addresses allocated for data storage. Far pointers are built from these data in routine ServiceMode6().

Signed integers:

darkcurrent is the value obtained in various routines which represents stray light and amplifier offsets in the absence of laser illumination. It is used in the routine reduce() (found in reduce.c) where it is subtracted from the measured DC intensity. It is imperative that the number of samples acquired match that of the data, as reduce() does no correction for sample size.

Unsigned integers:

counter1 is the 16 bit value to be loaded into counter 1 of the 8254 Programmable Interval Timer on the DASH-16 board. Allowed values are between 2 and 65536.

counter2 ditto for counter 2.

Unsigned long integers:

cnttime holds the collection time interval, in milliseconds. It is defined and modified within routine changetime() in module timedacq.c. It is used by timedacq(), changetime(), and gettdata().

Long integers:

begin_secs is the number of seconds since 1980 at the time we start a data collection run. It is stored in the output file so that the user can reconstruct the time and date the run was taken.

end_secs ditto for end of run.

total_secs is the difference between the above, hence, the elapsed time of the run, in seconds.

Floats:

scale[5] contains the gain settings for DC and the four lock-in amplifiers. The user can change gain settings at will to maintain a good signal-to-noise ratio; the program automatically applies appropriate correction factors when reducing the data.

Double precision floats:

calOmega is the calibration value of the fundamental frequency, from which the ellipsometric parameter Δ is determined. calOmega is defined in calibrat() and used in reduce().

cal2Omega ditto for the second harmonic, from which Ψ is derived.

Characters:

origdir[MAXPATH] contains the path specification from which the program was launched. When the program terminates, we are dropped back in the directory from whence we came.

5) Static Variable Declarations

in the module file_m.c:

stordir[MAXPATH] is the path spec to which output files are written. It is defined as equal to origdir, but may be changed in routine setpath().

file[MAXFILE+MAXEXT] is the name of the output file. It is constructed from the date as

ddmmmyy.xxx where

dd=day

mmm=month

yy=year

xxx=a cardinal number

example: 04JUL86.017 is the 17th run taken on July 4, 1986.

comment[MAXLENGTH] is the 64 character comment string allowed at the end of the run. This is for archival purposes; relevant notes can be stored here for storage with the data. Note the 64 character limit is defined as MAXLENGTH in the header file "gad.h".

in the module timedacq.c:

Booleans:

stop — is TRUE whenever Mode 6 operation is to be terminated. It can only be set by the interrupt handlers ServiceMode6() and ServiceMode7().

Unsigned integers:

blkcntr — is the current index of the array block[]. It is used to derive the location of data in high memory.

nPoints — is the number of points acquired. It is used throughout ellips.exe.

Characters:

message — contains the message string displayed in the upper right-hand corner of the screen and used in the routine reporterror() (in dast_m.c).

in the module plotdata.c:

Booleans:

graph_valid — is TRUE when the graph is already calculated and displayed. This tells the plotting routines not to redo everything.

Integers:

vminx — is the minimum x coordinate of the viewport. It is given in absolute screen coordinates.

vminy — ditto for y.

vmaxx — maximum x coordinate.

vmaxy	Have you got the hang of this yet?
vminor	is 80% of vmaxy. It is used to clip the viewport to the top 80% of the screen.
gminx	is the minimum x coordinate of the plot. It is given in absolute screen coordinates.
gminy	You know.
gmaxx	You know.
gmaxy	You know.
gdelx	is gmaxx-gminx, or the length of the x-axis.
gdely	ditto for y.
gzeroy	is the location of the y-axis zero. It is halfway down the y-axis.
numpltpts	is the number of points that can be plotted on a given x-axis. The minimum value is set to 8. The maximum value is determined by computing the resolution of the video monitor.
ginterval	is the spacing (in screen units) between plotted points along the x-axis. The maximum value is computed from the resolution of the video monitor and the minimum is 1.
multiplier	is the number of points between pixels. For example, if we have 256 pixels along the x-axis, we can plot 256 points. If we have 700 points

to plot, we can only plot one out of every four. Thus 175 points are plotted: point 1, point 5, point 9, etc. In this example, 'multiplier'=4.

6) Module Descriptions

GAD.H is a header file which contains a few miscellaneous but vital lines. Here we #define the variable type boolean as an integer, and assign values to the constants TRUE and FALSE. Also, MAXLENGTH is defined as 64, the maximum number of characters allowed in the comment string. Lastly, a function prototype is defined, that of the DASH-16 call routine TCM_DASG. This call is for Turbo C, medium memory model (far code pointers, near data pointers), revision G of the PCF-16 software package sold by MetraByte.

ELLIPS.C All global variables are declared here, and many are assigned default values. The startup screen and control-break handler c_break() are here.

INIT_M.C This is the initialization routine. Several important things happen here:

1) the current path specification is saved, to be used at program exit

2) memory is allocated for data storage through a call to allocseg()

3) the DASH-16 is initialized

4) the channel multiplexer is initialized

5) the timer/counters are initialized

6) software interrupts are enabled

CALIBRAT.C The all-important calibration routines are here. The user is instructed on setup of the optics and data are taken. Not all the data are used for each sample:

1) dark current readings are taken. Only the DC datum is stored, to be subtracted out from subsequent data.

2) 2ω readings are then performed, using DC and real & imaginary 2ω signals. Data are dark current corrected, scale factor corrected, ratioed to DC, and vector added to give the value 'cal2Omega'.

3) ω readings are performed, ditto above.

Mode 4 is used to acquire data to give a slight increase in speed (the timer is set to 40 kHz). 10,000 samples are acquired to give 1% accuracy.

The routine getscale() polls the lock-in amplifiers to get the channel gain settings. Due to the hardware configuration of the latches used (74LS367), this routine is pretty weird:

1) scale[] is initialized to 1

2) the four digital output lines are loaded in turn (0001, 0010, 0100, 1000) and written out to the latches on the lock-ins.

3) After each write (mode 13), the digital input lines are read via mode 14.

4) 0000 is sent to the output lines, shutting them off. The bit fields (stored in switches[]) are stored as follows:

switches[0] contains

bit 3 - im Omega post 20

bit 2 - re Omega post 20

bit 1 - re Omega pre 40

bit 0 - re Omega pre 20

switches[1] contains

bit 2 - im Omega pre 40

bit 1 - im Omega pre 20

switches[2] contains

bit 3 - im 2Omega post 20

bit 2 - re 2Omega post 20

bit 1 - re 2Omega pre 40

bit 0 - re 2Omega pre 20

switches[3] contains

bit 2 - im 2Omega pre 40

bit 1 - im 2Omega pre 20

5) The bit fields are deconvoluted.

6) The scale[] array is set by reading the bits in switches[].

ACQUIRE.C contains the screen which selects acquisition mode.

MANUAL.C This routine permits manual data acquisition. manual() is the screen which oversees this operation. getmdata() is the actual data-taking part of the module. Data are gathered via mode 3 and stored in high memory, reduced and plotted. Note that each point is actually the sum of 16 readings. This is to give a better signal-to-noise ratio. The discrepancy between this and the number of calibration readings is unimportant, as readings are all ratioed to their DC values (normalized). See REDUCE.C for more on this.

TIMEDACQ.C is designed for fast data! The maximum throughput has been limited to 1 millisecond per point, or about 6000 readings per second. At this rate, DMA must be employed. Interrupts are required to keep an eye on things, and several routines have been written to provide these services.

The routine timedacq() is the screen which controls this mode. Among the routines called are calibrat(), changetime(), gettdata(), and getfdata(). changetime() permits the user to change the time between points, relying on the routines calcdivisors() and loadtimer() in module DAST_M.C to do the work.

gettdata() is used to acquire data at intervals greater than 56 milliseconds. This is the amount of time needed by the AT&T 6300 to

reduce and plot one data set. gettdata() uses mode 4 to average 16 readings for each point.

getfdata() is the routine which does the fast acquisition. Two interrupt service routines (ISRs) are loaded to keep the program moving at a fast rate:

ServiceMode6() replaces the interrupt called when the DASH-16 reaches terminal count. The new routine starts the next segment of data-taking and reinstalls itself, as the call to mode 6 screws up the interrupt vector. If we run out of memory, the routine is supposed to exit gracefully.

ServiceMode7() catches the keyboard interrupt. Any keystroke will initiate this interrupt, which is used to cancel acquisition. There is a bogus call within the ISR (reply=getch()) to clear the keystroke from the keyboard buffer. This is necessary to avoid conflict with input of the comment string.

Install_DASH_ISR() is the first of the interrupt-related routines. It gets the interrupt vector from low memory and replaces it with a pointer to our interrupt. The old pointer is saved.

Remove_DASH_ISR() restores the old interrupt vector.

New_DASH_ISR() is the target of the newly installed interrupt vector. It first executes the old routine (to make sure any machine-related

functions receive attention) and then invokes the new routine, which in our case is ServiceMode6().

Ditto for the other routines, which handle the keyboard interrupt functions.

REDUCE.C Here is the data reduction routine. It is straightforward and well commented, in my humble opinion.

Things are complicated by the presence of the functions boxcar() and finish(). These guys provide the 21-point smooth. Read 'em and weep.

FILE_M.C getpath() uses createfile() and getdirectory() to find the highest-numbered occurrence of a date-matched file name via the routine findfile(). It then redirects the path spec to the directory containing that file. The mechanism behind this routine can be found in the text 'Turbo C DOS Utilities' by Robert Alonzo, Wiley, New York, 1988.

setpath() allows the user to change the path spec. The only thing which is a bit unusual is the use of Turbo C library functions fnsplit() and fnmerge() to assure a valid path spec.

storfile() writes the output file. The comment string is requested, parsed, and padded with ASCII 124 so that no blanks exist in the string. The file format is as follows:

filename.ext<cr>

comment<cr>

begin_secs<tab>end_secs<tab>counter1<tab>counter2<tab>

cnttime<tab>extchan<tab>acqmode<tab>calOmega<tab>

cal2Omega<cr>

nPoints<cr>

Delta<tab>Psi[<tab>extvalue]<cr>

.

.

.

where extvalue is an option dictated by the value of extchan.

loadfile() loads the above file format back in. The comment string is parsed and ASCII 124 is converted back into a space.

PLOTDATA.C These routines perform the rather crude plotting functions. If you want to make it pretty, go ahead. plotdata() is called with two parameters: the first tells the function to turn the plot on or off. The second is the number of points to plot. The routine is very simple. If a plot has not been defined (graph_valid is FALSE) and we want to plot something, then initialize the graphics and make the necessary screen calculations through sizedisplay(). If graph_valid is TRUE, plot the points or shut down the plot.

graphinit() looks for graphics hardware and loads the required files from disk. THIS IS IMPORTANT! The Turbo C graphics drivers MUST be resident in the same directory as the program, or else the program will halt.

sizedisplay() sets viewport coordinates for the screen resolution and scales the x- and y-axes. The x-axis is defined in multiples of 2 so that autoscaling is easy. That is, for a screen with x resolution of 300, the axis will be a maximum of 270, and 256 is the largest multiple of 2 that will fit.

drawaxis() draws and labels the axes.

graphpts() is the routine which places points on the screen. Three important variables are needed here. One is ginterval, the graphing interval between points. Another is multiplier, the number of points between pixels. Both are dependent on numpltpts, which is the number of plottable points on a given graph. numpltpts is the multiple of 2 which is equal to or greater than nPoints, e.g. if nPoints is 309, numpltpts is 512. If the screen resolution is 300 (see above), we can't accommodate all 309 points directly. ginterval would be 1, that is, every pixel would contain a point, but the multiplier would be 2, and every other point would be plotted. Hence points 1,3,5,...,309 would be plotted on an x-axis of 512, filling 155 pixels of the 256.

If numpltpts is insufficient to plot all points (numpltpts < nPoints) then a rescale routine is invoked and the plot is redrawn.

shutdown() turns off the plot and restores the text mode.

printtext() places a string in the lower viewport. There is room for two lines of text.

DAST_M.C This routine contains many of the functions and utilities used throughout the program. It's kind of a library.

beep() makes a beep.

intptr() constructs a pointer to high memory for placement and retrieval of data. It is optimized for 12 byte data, enough for 6 samples, thus one point.

Ditto for floatptr(), except type float.

reporterror() notifies the user of a fatal error in programming of the DASH-16 board. *It* is assumed that this would corrupt the run, so the routine forces the user to quit the program. It restores the old interrupt vectors, too.

writelabel() places a string in the upper right-hand corner of the text window. This lets the user know where he is in the program.

allocseg() sets up a table of segment addresses in the variable block[] and allocates the memory in the call defining block[0]. This is a dummy pointer; a loop is set up to assign the proper pointer values.

calcdivisors() is a routine based on an algorithm I designed in high school. It seeks a combination of counter1 and counter2 that provides a desired clock rate for data conversion. The routine computes the required divisor and then takes the square root. This is the geometric

mean of the divisor. The root is then decremented to 2, which is the minimum value for the counters. An error term is stored (along with associated value) and the values which produce minimum error between the desired and achievable values are returned.

loadtimer() loads the timer counters. The values counter1 and counter2 are unsigned integers for ease of computation in the routine calcdivisors(), but must be converted to two's complement form for programming the DASH-16.

7) Known Bugs

I am aware of three bugs in the program. The first two are annoyances, but do not corrupt anything, so I haven't bothered to correct them. The third occurs at the DOS level, and I haven't a clue as to a fix.

Bug One: File naming. If the program is running when the date wraps past midnight, the file name is not updated. That is, instead of creating a new file name based on the new date and restarting the numbering sequence from .001, the run is labeled according to the sequence for the old date. It probably isn't a difficult fix, but why bother?

Bug two: For some unknown reason, when calling the routine changetime(), the new time echoed back to the user is not displayed correctly. Example: 100 milliseconds is displayed as 10 milliseconds. I don't know why, and since the correct values are computed and loaded into the counters, why bother?

Bug three: Occasionally, the error message "No graphics hardware detected!" comes up. When this happens, the run is lost. The program aborts and returns control to DOS, from where it can be restarted. This bug may well be the result of cosmic rays, for it is rare and is totally unpredictable.

Curriculum Vitae

The author was born on March 10, 1958 in McKeesport, Pennsylvania. He graduated in 1980 with a Bachelor of Arts in Biophysics from The Johns Hopkins University, Baltimore, Maryland. Following graduation he joined the technical staff at the National Bureau of Standards, where he studied the corrosion behavior of surgical implant materials for the Metallurgy Division. His professional background includes extensive experience in analytical transmission electron microscopy, fracture mechanics of large steel plates, and synchrotron x-ray studies of iron-chromium alloys. He has also worked in the private sector, managing testing and field installation of Electron Energy Loss Spectrometers for Gatan, Inc. He had studied at the Masters' level in the Engineering Materials program at the University of Maryland, College Park, Maryland, before matriculating in the graduate program of the Department of Materials Science and Engineering at The Johns Hopkins University, Baltimore, Maryland. He is an associate member of the Johns Hopkins chapter of Sigma Xi, and a member of the National Association of Corrosion Engineers, the Chesapeake Society for Electron Microscopy, and the Electron Microscopy Society of America.